

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

10/006780

L6 FILE 'REGISTRY' ENTERED AT 14:27:22 ON 07 OCT 2003
1217 S ATPASE?/CN

- key terms

L6 FILE 'HCAPLUS' ENTERED AT 14:27:30 ON 07 OCT 2003
L6 1217 SEA FILE=REGISTRY ABB=ON PLU=ON ATPASE?/CN
L7 166 SEA FILE=HCAPLUS ABB=ON PLU=ON (PLASMODIUM OR FALCIPARU
M) AND (L6 OR ATPASE OR ADENOSINE(2W)(TRIPHOSPHATASE OR
TRI PHOSPHATASE))
L8 4 SEA FILE=HCAPLUS ABB=ON PLU=ON L7 AND (MICROTUB? OR
MICRO TUBUL?)

L6 1217 SEA FILE=REGISTRY ABB=ON PLU=ON ATPASE?/CN
L7 166 SEA FILE=HCAPLUS ABB=ON PLU=ON (PLASMODIUM OR FALCIPARU
M) AND (L6 OR ATPASE OR ADENOSINE(2W)(TRIPHOSPHATASE OR
TRI PHOSPHATASE))
L9 87 SEA FILE=HCAPLUS ABB=ON PLU=ON L7 AND (PROTEIN OR
PEPTIDE OR POLYPROTEIN OR POLYPEPTIDE)
L10 9 SEA FILE=HCAPLUS ABB=ON PLU=ON L9 AND STIMUL?

L12 11 L8 OR L10

L12 ANSWER 1 OF 11 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2003:454501 HCAPLUS

DOCUMENT NUMBER: 139:35072

TITLE: Vectors comprising nucleotide sequences for
target immunogen, PI31, CIIIA and antisense
HERNA mRNA, and their uses including use as
vaccines

INVENTOR(S): McCreavy, David Thomas; Fraser, William Duncan;
Gallagher, James Anthony

PATENT ASSIGNEE(S): University of Liverpool, UK

SOURCE: PCT Int. Appl., 52 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2003048371	A2	20030612	WO 2002-GB5512	20021206
WO 2003048371	A3	20030912		
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG			

PRIORITY APPLN. INFO.: GB 2001-29338 A 20011207

GB 2002-23829 A 20021012

AB The invention provides vectors (such as viral vectors, plasmid

THIS PAGE BLANK (USPTO)

vectors or phagemids) comprising: (a) a heterologous nucleotide sequence encoding an antigenic **polypeptide** from a pathogen (such as viral, bacterial, parasitic or fungal); (b) a nucleotide sequence encoding a protease inhibitor (such as human PI31); a nucleotide sequence for a constitutive, regulatable, and/or cell/tissue-specific promoter; and (d) a nucleotide sequence encoding an inhibitory RNA mols., specifically an antisense human HERNA oligonucleotides. The invention also provides vectors comprising a nucleotide sequence encoding CIITA, a **polypeptide** that **stimulates** the expression of MHC class II genes. The invention further provides the use of said vectors as vaccines in production of an immune response (humoral) to said antigens in an animal, such as human, wherein said vaccination may be against a viral, fungal, bacterial or parasitic disorder. Still further, the invention relates: (a) using said vectors in production of antibodies, wherein said antibodies may be of therapeutic and/or of diagnostic use; (b) that said vectors may be adapted for expression of humanized or chimeric antibodies; and (c) that said vectors may be used to used to immunize animals for production of hybridomas expressing a monoclonal antibody against antigen of interest. Finally, the invention provides the cDNA sequences of mouse CIITA, and human PI31; and partial cDNA sequence of human HERNA helicase. The invention related that the use of said vectors containing said sequences can be used to enhance secretion of translated immunogen, and enhance DNA vaccination bias away from an MHC class I event towards MHC class II event. In the examples, the invention presented the construction of two vectors, pCDNAFinal and pCDNA6TR-IRES-CIITA, wherein pCDNAFinal contains nucleotide sequences encoding immunogen parathyroid hormone-related **protein** (PTHrP), antisense HERNA mRNA, PI31 and CD4+ T-cell epitope from lymphocytic choriomeningitis virus and wherein pCDNA6TR-IRES-CIITA encodes CIITA. Specifically, the invention related that: (a) antisense HERNA RNA can increase the transcriptional efficiency of vectors resulting in greater levels of transgene expression; (b) PI31 can inhibit proteasome digestion of recombinant antigen making it more assessable to MHC class II antigens; (c) inclusion of CD4+ T-cell epitope ensured that degraded immunogen-MHC class II complex bound to CD4+ T cells; and (d) inclusion of CIITA **protein** allowed for over-expression of MHC class II antigens.

IT 9000-83-3, **Atpase**

RL: BSU (Biological study, unclassified); BIOL (Biological study) (promoter of; vectors comprising sequences for promoter, target immunogen, PI31, CIITA and antisense HERNA mRNA, and their uses, including as vaccines towards various disorders)

L12 ANSWER 2 OF 11 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2003:454460 HCAPLUS

DOCUMENT NUMBER: 139:31820

TITLE: **Plasmodium falciparum**
kinesin motor **protein** KinI-1 with
microtubule-stimulated
ATPase activity and uses for diagnosis
and treatment of malaria

INVENTOR(S): Sakowicz, Roman; Beraud, Christophe; Guo, Jun;
Freedman, Richard

PATENT ASSIGNEE(S): Cytokinetics, Inc., USA

SOURCE: PCT Int. Appl., 68 pp.

THIS PAGE BLANK (USPTO)

10/006780

CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 2
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2003048320	A2	20030612	WO 2002-US38360	20021127
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG				
US 2003104496	A1	20030605	US 2001-6780	20011130
PRIORITY APPLN. INFO.:			US 2001-6780	A2 20011130
			US 2002-86935	A2 20020228
AB The present invention relates to protein and cDNA sequences of a new kinesin motor protein , P. falciparum KinI-1 (PfKinI-1), and the use of these comps. for the diagnosis, treatment, or prevention of malaria. The invention provides isolated nucleic acid and amino acid sequences of kinesin superfamily motor protein KinI-1, which has microtubule stimulated ATPase activity and/or depolymerizes microtubules . The invention further relates to antibodies to PfKinI-1, methods of screening for PfKinI-1 modulators using biol. active PfKinI-1, and kits for screening for PfKinI-1 modulators.				
IT 9000-83-3, ATPase RL: BSU (Biological study, unclassified); BIOL (Biological study) (microtubule-stimulated; Plasmodium falciparum kinesin motor protein KinI-1 with microtubule-stimulated ATPase activity and uses for diagnosis and treatment of malaria)				
L12 ANSWER 3 OF 11 HCAPLUS COPYRIGHT 2003 ACS on STN ACCESSION NUMBER: 2003:435206 HCAPLUS DOCUMENT NUMBER: 139:909 TITLE: (Sequences of Plasmodium falciparum kinesin KinI-1 and use for treating malaria)				
INVENTOR(S): Sakowicz, Roman; Beraud, Christophe; Guo, Jun; Freedman, Richard				
PATENT ASSIGNEE(S): Cytokinetics, Inc., USA				
SOURCE: U.S. Pat. Appl. Publ., 32 pp. CODEN: USXXCO				
DOCUMENT TYPE: Patent				
LANGUAGE: English				
FAMILY ACC. NUM. COUNT: 2				
PATENT INFORMATION:				

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
------------	------	------	-----------------	------

Searcher : Shears 308-4994

THIS PAGE BLANK (USPTO)

US 2003104496	A1	20030605	US 2001-6780	20011130
WO 2003048320	A2	20030612	WO 2002-US38360	20021127

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG

PRIORITY APPLN. INFO.: US 2001-6780 A2 20011130
US 2002-86935 A2 20020228

AB The invention provides sequences of **Plasmodium falciparum** kinesin KinI-1. The invention also relates to the use of KinI-1 for the diagnosis, treatment, or prevention of malaria.

IT 9000-83-3, ATPase

RL: BSU (Biological study, unclassified); BIOL (Biological study) (microtubule-stimulated; sequences of **Plasmodium falciparum** kinesin KinI-1 and use for treating malaria)

L12 ANSWER 4 OF 11 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2002:752115 HCAPLUS

DOCUMENT NUMBER: 137:289734

TITLE: Sequence of **Plasmodium falciparum** chromosomes 2, 10, 11 and 14

AUTHOR(S): Gardner, Malcolm J.; Shallom, Shamira J.; Carlton, Jane M.; Salzberg, Steven L.; Nene, Vishvanath; Shoaibi, Azadeh; Ciecko, Anne; Lynn, Jeffery; Rizzo, Michael; Weaver, Bruce; Jarrahi, Behnam; Brenner, Michael; Parvizi, Babak; Tallon, Luke; Moazzez, Azita; Granger, David; Fujii, Claire; Hansen, Cheryl; Pederson, James; Feldblyum, Tamara; Peterson, Jeremy; Suh, Bernard; Angiuoli, Sam; Perte, Mihaela; Allen, Jonathan; Selengut, Jeremy; White, Owen; Cummings, Leda M.; Smith, Hamilton O.; Adams, Mark D.; Venter, J. Craig; Carucci, Daniel J.; Hoffman, Stephen L.; Fraser, Claire M.

CORPORATE SOURCE: The Institute for Genomic Research, Rockville, MD, 20850, USA

SOURCE: Nature (London, United Kingdom) (2002), 419(6906), 531-534
CODEN: NATUAS; ISSN: 0028-0836

PUBLISHER: Nature Publishing Group

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The mosquito-borne malaria parasite **Plasmodium falciparum** kills an estimated 0.7-2.7 million people every year, primarily children in sub-Saharan Africa. Without effective interventions, a variety of factors-including the spread of parasites resistant to antimalarial drugs and the increasing insecticide resistance of mosquitoes-may cause the number of malaria

THIS PAGE BLANK (USPTO)

cases to double over the next two decades. To **stimulate** basic research and facilitate the development of new drugs and vaccines, the genome of **Plasmodium falciparum** clone 3D7 has been sequenced using a chromosome-by-chromosome shotgun strategy. This report describes nucleotide sequences of chromosomes 10, 11 and 14, and a re-anal. of the chromosome 2 sequence. These chromosomes represent about 35% of the 23-megabase **P. falciparum** genome. The sequences are deposited in GenBank/EMBL/DDBJ under accession nos. AE001362.2 (chromosome 2), AE014185 (chromosome 10), AE014186 (chromosome 11), and AE014187 (chromosome 14).

only NA
no other

IT 465605-54-3

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)

(amino acid sequence; complete sequence of **Plasmodium falciparum** chromosomes 2, 10, 11 and 14)

REFERENCE COUNT: 30 THERE ARE 30 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L12 ANSWER 5 OF 11 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2002:485238 HCAPLUS

DOCUMENT NUMBER: 137:242895

TITLE: Comparative genomic analysis in the region of a major **Plasmodium**-refractoriness locus of *Anopheles gambiae*

AUTHOR(S): Thomasova, Dana; Ton, Lucas Q.; Copley, Richard R.; Zdobnov, Evgeny M.; Wang, Xuelan; Hong, Young S.; Sim, Cheolho; Bork, Peer; Kafatos, Fotis C.; Collins, Frank H.

CORPORATE SOURCE: European Molecular Biology Laboratory, Heidelberg, 69117, Germany

SOURCE: Proceedings of the National Academy of Sciences of the United States of America (2002), 99(12), 8179-8184

CODEN: PNASA6; ISSN: 0027-8424

PUBLISHER: National Academy of Sciences

DOCUMENT TYPE: Journal

LANGUAGE: English

AB We have sequenced six overlapping clones from a library of bacterial artificial chromosome (BAC) clones derived from a laboratory strain of the mosquito, *Anopheles gambiae*, the major vector of human malaria in Africa. The resulting uninterrupted 528-kb sequence is from the 8C region of the mosquito 2R chromosome, at or very near the major refractoriness locus associated with melanotic encapsulation of parasites. This sequence represents the first extensive view of the mosquito genome structure encompassing 48 genes. Genomic comparison reveals that the majority of the orthologues are found in six microsyntenic clusters in *Drosophila melanogaster*. A BAC clone that is wholly contained within this region demonstrates the existence of a remarkable degree of local polymorphism in this species, which may prove important for its population structure and vectorial capacity.

IT 9000-83-3

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)

(V-type, sequence homolog to; comparative genomic anal. in the region of a major **Plasmodium**-refractoriness locus of *Anopheles gambiae*)

THIS PAGE BLANK (USPTO)

10/006780

REFERENCE COUNT: 40 THERE ARE 40 CITED REFERENCES AVAILABLE
FOR THIS RECORD. ALL CITATIONS AVAILABLE
IN THE RE FORMAT

L12 ANSWER 6 OF 11 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2001:394971 HCAPLUS

DOCUMENT NUMBER: 135:118715

TITLE: Expression and functional characterization of a
Plasmodium falciparum Ca²⁺-
ATPase (PfATP4) belonging to a subclass
unique to apicomplexan organisms

AUTHOR(S): Krishna, Sanjeev; Woodrow, Charles; Webb,
Richard; Penny, Jeff; Takeyasu, Kunio; Kimura,
Masatsugu; East, J. Malcolm

CORPORATE SOURCE: Department of Infectious Diseases, St. George's
Hospital Medical School, London, SW17 ORE, UK

SOURCE: Journal of Biological Chemistry (2001), 276(14),
10782-10787

CODEN: JBCHA3; ISSN: 0021-9258

PUBLISHER: American Society for Biochemistry and Molecular
Biology

DOCUMENT TYPE: Journal

LANGUAGE: English

AB We have obtained a full-length P type **ATPase** sequence
(PfATP4) encoded by **Plasmodium falciparum** and
expressed PfATP4 in *Xenopus laevis* oocytes to study its function.
Comparison of the hitherto incomplete open reading frame with other
Ca²⁺-**ATPase** sequences reveals that PfATP4 differs
significantly from previously defined categories. The
Ca²⁺-dependent **ATPase** activity of PfATP4 is
stimulated by a much broader range of [Ca²⁺]_{free} (3.2-320
μM) than are an avian SERCA1 pump or rabbit SERCA 1a (maximal
activity < 10 μM). The activity of PfATP4 is resistant to
inhibition by ouabain (200 μM) or thapsigargin (0.8 μM) but is
inhibited by vanadate (1 mM) or cyclopiazonic acid (1 μM). We
used a quant. polymerase chain reaction to assay expression of mRNA
encoding PfATP4 relative to that for β-tubulin in synchronized
asexual stages and found variable expression throughout the life
cycle with a maximal 5-fold increase in meronts compared with ring
stages. This anal. suggests that PfATP4 defines a novel subclass of
Ca²⁺-**ATPases** unique to apicomplexan organisms and
therefore offers potential as a drug target.

IT 9000-83-3, **ATPase**

RL: BAC (Biological activity or effector, except adverse); BPR
(Biological process); BSU (Biological study, unclassified); PRP
(Properties); BIOL (Biological study); PROC (Process)
(calcium-activated, PfATP4; expression and functional
characterization of a **Plasmodium falciparum**
Ca²⁺-**ATPase** (PfATP4) belonging to a subclass unique to
apicomplexan organisms)

REFERENCE COUNT: 38 THERE ARE 38 CITED REFERENCES AVAILABLE
FOR THIS RECORD. ALL CITATIONS AVAILABLE
IN THE RE FORMAT

L12 ANSWER 7 OF 11 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1994:694055 HCAPLUS

DOCUMENT NUMBER: 121:294055

TITLE: **Plasmodium falciparum**:

THIS PAGE BLANK (USPTO)

further characterization of putative cation
ATPases
 AUTHOR(S): Krishna, Sanjeev; Cowan, Gill M.; Robson,
 Kathryn J.; Meade, John C.
 CORPORATE SOURCE: Inst. Mol. Med., John Radcliffe Hosp., Oxford,
 OX3 9DU, UK
 SOURCE: Experimental Parasitology (1994), 78(1), 113-17
 CODEN: EXPAAA; ISSN: 0014-4894
 DOCUMENT TYPE: Journal
 LANGUAGE: English
 AB The emergence of multi-drug resistant strains of **Plasmodium**
falciparum and the current lack of a vaccine have
stimulated research to identify new chemotherapeutic
 targets. The authors have focused on the isolation and
 characterization of nucleotide sequences encoding putative "P"-type
 cation **ATPases** from **P. falciparum**, because
 selective inhibitors of mammalian members of this transporter family
 are well studied, and some are already in clin. use. This report
 describes the isolation and sequence anal. of two addnl. members of
 this family, **ATPases** 2 and 3, and their differential
 expressivity during the erythrocytic stage of parasite development.
 Amino acid sequence anal. identified the 7 amino acid motifs present
 at the phosphorylation site (DKTGTLT) of other P-type cation
ATPases, in both clones. **ATPase** 2 was localized
 to chromosome 5 and **ATPase** 3 to chromosome 12 on a
 Southern blot of chromosomes resolved by pulsed-field gel
 electrophoresis. Northern blot anal. showed that both
ATPases 2 and 3 were expressed during the erythrocytic stage
 of the infection. **ATPase** 2 has a mRNA transcript size of
 5.5-6 kb and is expressed at the schizont stage of development.
ATPase 3 mRNA is approx. 5 kb in size and the strongest
 signal corresponds to mRNA from parasites at the ring stage of
 development (12-24 h postinvasion), although mRNA is also detected
 in all other erythrocytic stages examined
 IT 9000-83-3, **ATPase**
 RL: PRP (Properties)
 (2 and 3; sequence and chromosomal localization and developmental
 expression of genes encoding P-type cation **ATPases** 2
 and 3 of **Plasmodium falciparum**)
 L12 ANSWER 8 OF 11 HCAPLUS COPYRIGHT 2003 ACS on STN
 ACCESSION NUMBER: 1994:407323 HCAPLUS
 DOCUMENT NUMBER: 121:7323
 TITLE: Immunomodulatory **peptides** binding to
 human major histocompatibility complex (MHC)
 class II allotype
 INVENTOR(S): Urban, Robert Glen; Chicz, Roman M.; Vignali,
 Dario A. A.; Hedley, Mary Lynne; Stern, Lawrence
 J.; Strominger, Jack L.
 PATENT ASSIGNEE(S): President and Fellows of Harvard College, USA
 SOURCE: PCT Int. Appl., 59 pp.
 CODEN: PIXXD2
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 1
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
------------	------	------	-----------------	------

Searcher :	Shears	308-4994
------------	--------	----------

THIS PAGE BLANK (USPTO)

-----	-----	-----	-----	-----
WO 9404557	A1	19940303	WO 1992-US6692	19920811
W: JP				
JP 08502244	T2	19960312	JP 1992-506181	19920811
PRIORITY APPLN. INFO.:			WO 1992-US6692	19920811

AB A purified oligopeptide preparation comprising an amino acid sequence identical to that of a segment of a naturally-occurring human **protein** that binds to human major histocompatibility complex (MHC) class II allotype is provided. The human **protein** is an MHC class I or II mol., HLA-A2, invariant chain (Ii), etc.,. A method is described for inhibiting an immune response in a human patient by contacting an antigen-presenting cell (APC) of the patient with a therapeutic composition or an immune-**stimulating** complex (ISCOM) containing the oligopeptide, or by expression of the oligopeptide-coding sequence linked to a trafficking sequence in APCs. The oligopeptide also can be used for inducing an immune response against pathogens. The options of the oligopeptide delivery system is also described. Purification and characterization of 6 HLA-DR antigens (HLA-DR1.apprx.4; HLA-DR7.apprx.8) from Epstein-Barr virus-transformed human B lymphoblastoid cell lines were demonstrated.

IT **9000-83-3, ATPase**
 RL: BIOL (Biological study)
 (Na+/K+, **peptides** of, as immunomodulators)

L12 ANSWER 9 OF 11 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1991:673744 HCAPLUS

DOCUMENT NUMBER: 115:273744

TITLE: **Stimulation** of the interaction between actin and myosin by Physarum caldesmon-like **protein** and smooth muscle caldesmon

AUTHOR(S): Ishikawa, Ryoki; Okagaki, Tsuyoshi; Higashi-Fujime, Sugie; Kohama, Kazuhiro

CORPORATE SOURCE: Sch. Med., Gunma Univ., Maebashi, 371, Japan

SOURCE: Journal of Biological Chemistry (1991), 266(32), 21784-90

CODEN: JBCHA3; ISSN: 0021-9258

DOCUMENT TYPE: Journal

LANGUAGE: English

AB An actin-binding **protein** was purified from the **plasmodia** of a lower eukaryote, *P. polycephalum*, with an apparent mol. weight of 210 kDa on SDS-PAGE. This **protein** bound to actin filaments with a stoichiometry of 1:7-8 in a Ca²⁺-calmodulin-dependent manner. Antibody raised against caldesmon from smooth muscle cross-reacted with the 210-kDa **protein**. In vitro motility assay revealed that the 210-kDa **protein** increased the sliding velocity of actin filaments on Physarum myosin. The 210-kDa **protein** more than doubled the actin-activated **ATPase** activity of Physarum myosin under comparative conditions of in vitro motility assay. Further increases in the concentration of the 210-kDa **protein** decreased its **stimulatory** effects. Ca²⁺-calmodulin prevented the **stimulatory** effects of the 210-kDa **protein**. Unexpectedly, smooth muscle caldesmon also increased the sliding velocity of actin filaments on smooth muscle myosin at lower concns. The well-known inhibitory effect of smooth muscle caldesmon on the actin-myosin interaction was observed with this motility assay when the concentration of the caldesmon was increased further. The

THIS PAGE BLANK (USPTO)

stimulatory and inhibitory effects were confirmed by measurements of actin-activated **ATPase** activity of smooth muscle myosin. From estns. of the intracellular concns. of the 210-kDa **protein** and smooth muscle caldesmon in vivo, it appears that effects of the former and the latter on actin-myosin interactions in vivo are **stimulatory** and inhibitory, resp.

IT 9000-83-3, **ATPase**

RL: BIOL (Biological study)
(of myosin, actin-binding **protein** of Physarum polycephalum and caldesmon of smooth muscle effect on)

L12 ANSWER 10 OF 11 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1989:3408 HCAPLUS

DOCUMENT NUMBER: 110:3408

TITLE: Purification of **Plasmodium falciparum** digestive vacuoles and partial characterization of the vacuolar membrane **ATPase**

AUTHOR(S): Choi, Inpyo; Mego, John L.

CORPORATE SOURCE: Dep. Biol., Univ. Alabama, Tuscaloosa, AL, USA

SOURCE: Molecular and Biochemical Parasitology (1988), 31(1), 71-8

CODEN: MBIPDP; ISSN: 0166-6851

DOCUMENT TYPE: Journal

LANGUAGE: English

AB P. **falciparum** digestive vacuoles containing Fe³⁺ oxide granules were purified from parasite homogenates by centrifugation on discontinuous sucrose gradients. Digestive vacuole membranes prepared by osmotic lysis and washed with KCl showed no detectable contamination by erythrocyte membrane **proteins** and only minimal contamination by nonvacuolar parasite **proteins**. Purified vacuolar membranes were 2.6-fold enriched in total parasite membrane **ATPase** activity. This **ATPase** was optimally active at pH 7 in the presence of >2 mM Mg²⁺. Ca²⁺ and Mn²⁺ were .apprx.80-90% as effective as Mg²⁺, and Zn²⁺, Co²⁺, and Fe²⁺ also exerted some **stimulatory** effect. The vacuolar membrane also hydrolyzed GTP, UTP, CTP, and ADP, but AMP and 3',5'-cAMP were hydrolyzed only one-tenth as effectively as ATP. The **ATPase** was unaffected by vanadate, ouabain, or oligomycin but was significantly inhibited by the proton pump inhibitors NEM and NBD-Cl. Of 6 antimalarial drugs tested, quinine and quinacrine were the most effective inhibitors and mefloquine was the least effective.

IT 9000-83-3P, **ATPase**

RL: PREP (Preparation)
(of digestive vacuoles of **Plasmodium falciparum**, purification and characterization of)

L12 ANSWER 11 OF 11 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1982:434955 HCAPLUS

DOCUMENT NUMBER: 97:34955

TITLE: A novel 36,000-dalton actin-binding protein purified from microfilaments in Physarum **plasmodia** which aggregates actin filaments and blocks actin-myosin interaction

AUTHOR(S): Ogihara, Satoshi; Tonomura, Yuji

CORPORATE SOURCE: Fac. Sci., Osaka Univ., Osaka, 560, Japan

SOURCE: Journal of Cell Biology (1982), 93(3), 604-14

THIS PAGE BLANK (USPTO)

10/006780

CODEN: JCLBA3; ISSN: 0021-9525

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The **plasmodia** of *P. polycephalum* contains large aggregates of entangled actin microfilaments. Treatment with Triton X-100 gave a demembrated cytoskeleton consisting of entangled actin filaments which showed almost no interaction with rabbit skeletal myosin. A novel actin-binding protein purified from the cytoskeleton stoichiometrically binds to actin and causes actin filaments to curl and aggregate. The protein inhibits **ATPase** activity as well as the superpptn. of reconstituted rabbit skeletal muscle actomyosin. This protein has a mol. weight of 36,000 and binds 7 mol of actin/mol 36,000 polypeptide.

IT 9000-83-3

RL: BIOL (Biological study)
(inhibition of, of actomyosin, actin-binding protein inhibition of)

(FILE 'MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH, JICST-EPLUS, JAPIO' ENTERED AT 14:31:47 ON 07 OCT 2003)

L13 3 S L8

L14 21 S L10

L15 24 S L13 OR L14

L16 17 DUP REM L15 (7 DUPLICATES REMOVED)

L16 ANSWER 1 OF 17 WPIDS COPYRIGHT 2003 THOMSON DERWENT on STN

ACCESSION NUMBER: 2003-505298 [47] WPIDS

DOC. NO. CPI: C2003-135135

TITLE: New vector, useful for preparing a composition for treating or preventing bacterial, viral, fungal or parasitic infection.

DERWENT CLASS: B04 C06 D16

INVENTOR(S): FRASER, W D; GALLAGHER, J A; MCCREAVY, D T

PATENT ASSIGNEE(S): (UYLI-N) UNIV LIVERPOOL

COUNTRY COUNT: 101

PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 2003048371	A2	20030612	(200347)*	EN	52
RW: AT BE BG CH CY CZ DE DK EA EE ES FI FR GB GH GM GR IE IT KE					
LS LU MC MW MZ NL OA PT SD SE SI SK SL SZ TR TZ UG ZM ZW					
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ					
DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP					
KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ					
NO NZ OM PH PL PT RO RU SC SD SE SG SK SL TJ TM TN TR TT TZ					
UA UG US UZ VN YU ZA ZM ZW					

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 2003048371	A2	WO 2002-GB5512	20021206

PRIORITY APPLN. INFO: GB 2002-23829 20021012; GB 2001-29338 20011207

AN 2003-505298 [47] WPIDS

Searcher : Shears 308-4994

THIS PAGE BLANK (USPTO)

AB WO2003048371 A UPAB: 20030723

NOVELTY - An vector comprising a heterologous nucleic acid sequence encoding an antigenic **polypeptide** and a nucleic acid molecule comprising a 3188 base pair sequence, given in the specification, a nucleic acid molecule which hybridizes to it and which encodes a protease inhibitor **polypeptide**, or nucleic acid molecules which comprise degenerate nucleic acid sequences. The vector is adapted for the expression of each **polypeptide**.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

- (1) inducing an immune response to an antigenic **polypeptide**;
- (2) an antibody;
- (3) a cell transformed with the novel vector;
- (4) producing humanized or chimeric antibody;
- (5) a hybridoma cell line which produces a monoclonal antibody;
- (6) a vaccine comprising the novel vector; and
- (7) vaccinating an animal, preferably a human, against at least one pathological condition.

ACTIVITY - Antibacterial; Virucide; Antiparasitic; Antifungal; Anti-HIV; Antiulcer.

No biological data is given.

MECHANISM OF ACTION - Gene therapy; Vaccine.

USE - The vector is useful for preparing a composition for preventing or treating AIDS, herpes, rubeola, rubella, varicella, influenza, common cold or viral meningitis; septicemia, tuberculosis, bacteria-associated food poisoning, blood infections, peritonitis, endocarditis, sepsis, bacterial meningitis, pneumonia, stomach ulcers, gonorrhea, strep throat, streptococcal-associated toxic shock, necrotizing fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis, dysentery or shigellosis; Candidiasis; or trypanosomiasis, malaria, schistosomiasis or Chagas disease (claimed).

Dwg.0/9

L16 ANSWER 2 OF 17 EMBASE COPYRIGHT 2003 ELSEVIER INC. ALL RIGHTS RESERVED. on STN

ACCESSION NUMBER: 2003300235 EMBASE

TITLE: The pathophysiology of **falciparum** malaria.

AUTHOR: Clark I.A.; Cowden W.B.

CORPORATE SOURCE: I.A. Clark, Sch. of Biochem./Molecular Biology, Australian National University, Canberra, ACT 0200, Australia. ian.clark@anu.edu.au

SOURCE: Pharmacology and Therapeutics, (1 Aug 2003) 99/2 (221-260).

Refs: 466

ISSN: 0163-7258 CODEN: PTHDT

COUNTRY: United States

DOCUMENT TYPE: Journal; General Review

FILE SEGMENT: 004 Microbiology

017 Public Health, Social Medicine and Epidemiology

026 Immunology, Serology and Transplantation

037 Drug Literature Index

038 Adverse Reactions Titles

LANGUAGE: English

SUMMARY LANGUAGE: English

AB **Falciparum** malaria is a complex disease with no simple

THIS PAGE BLANK (USPTO)

explanation, affecting organs where the parasite is rare as well as those organs where it is more common. We continue to argue that it can best be understood in terms of excessive **stimulation** of normally useful pathways mediated by inflammatory cytokines, the prototype being tumor necrosis factor (TNF). These pathways involve downstream mediators, such as nitric oxide (NO) that the host normally uses to control parasites, but which, when uncontrolled, have bioenergetic failure of patient tissues as their predictable end point. **Falciparum** malaria is no different from many other infectious diseases that are clinically confused with it. The sequestration of parasitized red blood cells, prominent in some tissues but absent in others with equal functional loss, exacerbates, but does not change, these overriding principles. Recent opportunities to stain a wide range of tissues from African pediatric cases of **falciparum** malaria and sepsis for the inducible NO synthase (iNOS) and migration inhibitory factor (MIF) have strengthened these arguments considerably. The recent demonstration of bioenergetic failure in tissue removed from sepsis patients being able to predict a fatal outcome fulfils a prediction of these principles, and it is plausible that this will be demonstrable in severe **falciparum** malaria. Understanding the disease caused by **falciparum** malaria at a molecular level requires an appreciation of the universality of poly(ADP-ribose) polymerase-1 (PARP-1) and Na(+)/K(+)-ATPase and the protean effects of activation by inflammation of the former that include inactivation of the latter. .COPYRG. 2003 Elsevier Inc. All rights reserved.

L16 ANSWER 3 OF 17 EMBASE COPYRIGHT 2003 ELSEVIER INC. ALL RIGHTS RESERVED. on STN DUPLICATE 1

ACCESSION NUMBER: 2003191572 EMBASE

TITLE: Replication fork-**stimulated** eIF-4A from **Plasmodium** cynomolgi unwinds DNA in the 3' to 5' direction and is inhibited by DNA-interacting compounds.

AUTHOR: Tuteja R.; Tuteja N.; Malhotra P.; Chauhan V.S.

CORPORATE SOURCE: R. Tuteja, Intl. Ctr. for Genetic Eng./Biotech., Aruna Asaf Ali Marg, New Delhi 110067, India. renu@icgeb.res.in

SOURCE: Archives of Biochemistry and Biophysics, (1 Jun 2003) 414/1 (108-114).
Refs: 47
ISSN: 0003-9861 CODEN: ABBIA4

COUNTRY: United States

DOCUMENT TYPE: Journal; Article

FILE SEGMENT: 030 Pharmacology
037 Drug Literature Index

LANGUAGE: English

SUMMARY LANGUAGE: English

AB **Plasmodium** cynomolgi. DEAD-box DNA helicase 45 (PcDDH45) is an ATP-dependent DNA-unwinding enzyme with intrinsic DNA-dependent **ATPase** activity and is highly homologous to eIF-4A. In this study, we have further characterized and tested the effect of various DNA-interacting compounds on the DNA-unwinding activity of PcDDH45. The results show that PcDDH45 translocates in the 3' to 5' direction along the bound strand, a replication fork-like structure of the substrate **stimulates** its DNA-unwinding activity, and it failed to unwind blunt-ended duplex DNA. Of various compounds

THIS PAGE BLANK (USPTO)

tested, only cisplatin, 4',6'-diamidino-2-phenylindole, daunorubicin, and nogalamycin were inhibitory to the unwinding activity of PcDDH45 with apparent IC(50) values of 1.0, 4.0, 7.5, and 1.7 μ M, respectively. These results suggest that the interaction of these compounds with duplex DNA generate a complex that probably impedes the translocation of PcDDH45, resulting in inhibition of unwinding activity. This study is one of the first to demonstrate the effect of various DNA-binding compounds on a malaria parasite DNA helicase and should make an important contribution to our better understanding of the nucleic acid transactions in the parasite. .COPYRG. 2003 Elsevier Science (USA). All rights reserved.

L16 ANSWER 4 OF 17 SCISEARCH COPYRIGHT 2003 THOMSON ISI on STN
 ACCESSION NUMBER: 2002:547065 SCISEARCH
 THE GENUINE ARTICLE: 566XK
 TITLE: Multidrug resistance phenotype mediated by the P-glycoprotein-like transporter in Leishmania: A search for reversal agents
 AUTHOR: Perez-Victoria J M; Di Pietro A; Barron D; Ravelo A G; Castanys S; Gamarro F (Reprint)
 CORPORATE SOURCE: CSIC, Inst Parasitol & Biomed Lopez Neyra, C Ventanilla 11, Granada 18001, Spain (Reprint); CSIC, Inst Parasitol & Biomed Lopez Neyra, Granada 18001, Spain; CNRS, UMR 5086, Inst Biol & Chim Prot, Lyon, France; Univ Lyon 1, CNRS, Lab Prod Nat, F-69622 Villeurbanne, France; Univ La Laguna, Inst Bioorgan Antonio Gonzalez, Tenerife, Spain
 COUNTRY OF AUTHOR: Spain; France
 SOURCE: CURRENT DRUG TARGETS, (AUG 2002) Vol. 3, No. 4, pp. 311-333.
 Publisher: BENTHAM SCIENCE PUBL LTD, PO BOX 1673, 1200 BR HILVERSUM, NETHERLANDS.
 ISSN: 1389-4501.
 DOCUMENT TYPE: General Review; Journal
 LANGUAGE: English
 REFERENCE COUNT: 193

ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS

AB Protozoan parasites are responsible for important diseases that threaten the lives of nearly one-quarter of the human population world-wide. Among them, leishmaniasis has become the second cause of death, mainly due to the emergence of parasite resistance to conventional drugs. P-glycoprotein (Pgp)-like transporters overexpression is a very efficient mechanism to reduce the intracellular accumulation of many drugs in cancer cells and parasitic protozoans including *Plasmodium* and *Leishmania*, thus conferring a multidrug resistance (MDR) phenotype. Therefore, there is a great clinical interest in developing inhibitors of these transporters to overcome such a resistance. Pgps are active pumps belonging to the ATP-binding cassette (ABC) superfamily of **proteins**, and consist of two homologous halves, each containing a transmembrane domain (TMD) involved in drug efflux, and a cytosolic nucleotide-binding domain (NBD) responsible for ATP binding and hydrolysis. Most conventional cancer MDR modulators interact with the drug-binding sites on the TMDs of Pgps, but they are also usually transported and the required concentrations for a permanent inhibition produce subsequent side-effects that hamper their clinical use. Besides, they only poorly modulate the

THIS PAGE BLANK (USPTO)

resistance in protozoan parasites. We review here a rational strategy developed to overcome the MDR phenotype in *Leishmania*, consisting in: i) the selection of an MDR *Leishmania tropica* line that overexpresses a Pgp-like transporter; ii) the use of their cytosolic NBDs as new pharmacological targets; iii) the search of new natural compounds that revert the MDR phenotype in *Leishmania* by binding to the TMDs; iv) the combination of subdoses of the above selected modulators directed to both targets in the transporter, NBDs and TMDs, to accumulate their reversal effects while diminishing their toxicity. In this way, we have reverted the MDR phenotype in *Leishmania*, including the resistance to the most promising new antileishmania agents, the alkyl-lysophospholipids. This approach might be extrapolated to be used in other eukaryotic cells.

L16 ANSWER 5 OF 17 EMBASE COPYRIGHT 2003 ELSEVIER INC. ALL RIGHTS RESERVED. on STN

ACCESSION NUMBER: 2002097347 EMBASE
 TITLE: Cytoskeleton of apicomplexan parasites.
 AUTHOR: Morrisette N.S.; Sibley L.D.
 CORPORATE SOURCE: N.S. Morrisette, Department of Molecular Microbiology, Washington Univ. School of Medicine, 660 South Euclid Ave., St. Louis, MO 63110, United States. naomi@borcim.wustl.edu
 SOURCE: Microbiology and Molecular Biology Reviews, (2002) 66/1 (21-38).
 Refs: 200
 ISSN: 1092-2172 CODEN: MMBRF7
 COUNTRY: United States
 DOCUMENT TYPE: Journal; General Review
 FILE SEGMENT: 004 Microbiology
 030 Pharmacology
 037 Drug Literature Index
 LANGUAGE: English
 SUMMARY LANGUAGE: English

AB The Apicomplexa are a phylum of diverse obligate intracellular parasites including *Plasmodium* spp., the cause of malaria; *Toxoplasma gondii* and *Cryptosporidium parvum*, opportunistic pathogens of immunocompromised individuals; and *Eimeria* spp. and *Theileria* spp., parasites of considerable agricultural importance. These protozoan parasites share distinctive morphological features, cytoskeletal organization, and modes of replication, motility, and invasion. This review summarizes our current understanding of the cytoskeletal elements, the properties of cytoskeletal proteins, and the role of the cytoskeleton in polarity, motility, invasion, and replication. We discuss the unusual properties of actin and myosin in the Apicomplexa, the highly stereotyped **microtubule** populations in apicomplexans, and a network of recently discovered novel intermediate filament-like elements in these parasites.

L16 ANSWER 6 OF 17 MEDLINE on STN DUPLICATE 2
 ACCESSION NUMBER: 2001560947 MEDLINE
 DOCUMENT NUMBER: 21519010 PubMed ID: 11606229
 TITLE: **Microtubule** associated motor proteins of *Plasmodium falciparum* merozoites.
 AUTHOR: Fowler R E; Smith A M; Whitehorn J; Williams I T; Bannister L H; Mitchell G H
 CORPORATE SOURCE: Malaria Laboratory, Department of Immunobiology,

THIS PAGE BLANK (USPTO)

Guy's, King's and St Thomas' School of Medicine, KCL,
 Guy's Hospital, London Bridge, London, SE1 9RT, UK..
 ruth.fowler@ed.ac.uk

SOURCE: MOLECULAR AND BIOCHEMICAL PARASITOLOGY, (2001 Oct)
 117 (2) 187-200.
 Journal code: 8006324. ISSN: 0166-6851.

PUB. COUNTRY: Netherlands
 DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
 LANGUAGE: English
 FILE SEGMENT: Priority Journals
 ENTRY MONTH: 200201
 ENTRY DATE: Entered STN: 20011022
 Last Updated on STN: 20020201
 Entered Medline: 20020131

AB We have studied the occurrence, stage specificity and cellular location of key molecules associated with **microtubules** in **Plasmodium falciparum** merozoites. Antibodies to gamma tubulin, conventional kinesin and cytoplasmic dynein were used to determine the polarity of merozoite **microtubules** (mt), the stage specificity of the motor proteins and their location during merozoite development. We conclude that the minus ends of the mts are located at their apical pole. Kinesin was present throughout the lifecycle, appearing as a distinct crescent at the apex of developing merozoites. The vast majority of cytoplasmic dynein reactivity occurred in late merogony, also appearing at the merozoite apex. Destruction of mt with dinitroanilines did not affect the cellular location of kinesin or dynein. In invasion assays, dynein inhibitors reduced the number of ring stage parasites. Our results show that both conventional kinesin and cytoplasmic dynein are abundant, located at the negative pole of the merozoite mt and, intriguingly, appear there only in very late merogony, prior to merozoite release and invasion.

L16 ANSWER 7 OF 17 MEDLINE on STN

ACCESSION NUMBER: 2001297917 MEDLINE
 DOCUMENT NUMBER: 21273137 PubMed ID: 11378198
 TITLE: Two classes of plant-like vacuolar-type H(+)-pyrophosphatases in malaria parasites.
 AUTHOR: McIntosh M T; Drozdowicz Y M; Laroia K; Rea P A; Vaidya A B
 CORPORATE SOURCE: Department of Microbiology and Immunology, MCP Hahnemann University, 2900 Queen Lane, Philadelphia, PA 19129, USA.. michael.t.mcintosh@drexel.edu
 CONTRACT NUMBER: AF28398
 SOURCE: MOLECULAR AND BIOCHEMICAL PARASITOLOGY, (2001 May) 114 (2) 183-95.
 Journal code: 8006324. ISSN: 0166-6851.

PUB. COUNTRY: Netherlands
 DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
 LANGUAGE: English
 FILE SEGMENT: Priority Journals
 OTHER SOURCE: GENBANK-AF115766; GENBANK-AF115767; GENBANK-AF283528
 ENTRY MONTH: 200108
 ENTRY DATE: Entered STN: 20010806
 Last Updated on STN: 20010806
 Entered Medline: 20010802

AB In plants, cytosolic inorganic pyrophosphate (PP(i)) is hydrolyzed by energy-conserving vacuolar-type H(+)-pyrophosphatases (V-PPases)

THIS PAGE BLANK (USPTO)

that harness the free energy of PP(i) hydrolysis to establish transmembrane H(+) gradients. Here we describe the identification and cloning of two genes, PfVP1 and PfVP2, from the malaria parasite *Plasmodium falciparum*. Inferred to encode type I (K(+)-dependent) and type II (K(+)-independent) V-PPases, respectively, PfVP1 and PfVP2 appeared more sequence divergent from each other than from their type I and type II counterparts in plants. The steady state levels of PfVP1 mRNA were high in comparison to PfVP2 mRNA throughout the erythrocytic phases of infection. Western analyses of trophozoite membranes using generic V-PPase antibodies (PAB(HK) and PAB(TK)) demonstrated appreciable amounts of a Mr 67000 **polypeptide** whose associated aminomethylenediphosphonate- (AMDP) inhibitable PPase activity was markedly **stimulated** by K(+). Immunofluorescence microscopy of infected erythrocytes revealed PfVP antigen associated with both the parasite plasma membrane and punctate intracellular inclusions. Transient transfection of a PfVP1-GFP fusion further supported the localization of PfVP1 to the parasite plasma membrane. Based on these findings and the growth-retarding effects of AMDP, *P. falciparum* is concluded to possess both type I and type II V-PPases of which the former has the greatest potential for contributing to the establishment of H(+) gradients across the parasite plasma membrane under conditions of energy limitation.

L16 ANSWER 8 OF 17 SCISEARCH COPYRIGHT 2003 THOMSON ISI on STN
 ACCESSION NUMBER: 2001:373496 SCISEARCH
 THE GENUINE ARTICLE: 427QR
 TITLE: Detection and localization of a Ca²⁺-**ATPase** activity in *Toxoplasma gondii*
 AUTHOR: Bouchot A; Jaillet J D; Bonhomme A (Reprint); Pezzella-D'Alessandro N; Laquerriere P; Kilian L; Burlet H; Gomez-Marin J E; Pluot M; Bonhomme P; Pinon J M
 CORPORATE SOURCE: CHU Maison Blanche, IFR 53, UPRES EA 2070, 51 Rue Cognacq Jay, F-51095 Reims, France (Reprint); CHU Maison Blanche, IFR 53, UPRES EA 2070, F-51095 Reims, France; CHU Maison Blanche, IFR 53, Lab Parasitol Mycol, F-51095 Reims, France; Lab Microscopie Elect Analyt & Quantitat, IFR 53, Reims, France; Hop Robert Debre, Anat Pathol Lab, Reims, France; Univ Nacl Colombia, Hosp San Juan de Dios, Dept Med Interna, Grp Pathol Infecciosa, Santafe De Bogota, Colombia
 COUNTRY OF AUTHOR: France; Colombia
 SOURCE: CELL STRUCTURE AND FUNCTION, (FEB 2001) Vol. 26, No. 1, pp. 49-60.
 Publisher: JAPAN SOC CELL BIOLOGY, SHIMOTACHIURI OGAWA-HIGASHI, KAMIKYOKU KYOTO, 602, JAPAN.
 ISSN: 0386-7196.
 DOCUMENT TYPE: Article; Journal
 LANGUAGE: English
 REFERENCE COUNT: 59

ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS

AB *Toxoplasma gondii*, the agent causing toxoplasmosis, is an obligate intracellular protozoan parasite. A calcium signal appears to be essential for intracellular transduction during the active process of host cell invasion. We have looked for a Ca²⁺-transport **ATPase** in tachyzoites and found Ca²⁺-**ATPase**

THIS PAGE BLANK (USPTO)

activity (11-22 nmol Pi liberated/mg **protein**/min) in the tachyzoite membrane fraction, This ATP-dependent activity was **stimulated** by Ca²⁺ and Mg²⁺ ions and by calmodulin, and was inhibited by pump inhibitors (sodium orthovanadate or thapsigargin). We used cytochemistry and X-ray microanalysis of cerium phosphate precipitates and immunolabelling to find the Ca²⁺, Mg²⁺-**ATPase**. It was located mainly in the membrane complex, the conoid, nucleus, secretory organelles (rhoptries, dense granules) and in vesicles with a high calcium concentration, Thus, *Toxoplasma gondii* possesses Ca²⁺-pump **ATPase** (Ca²⁺, Mg²⁺-**ATPase**) as do eukaryotic cells.

L16 ANSWER 9 OF 17 SCISEARCH COPYRIGHT 2003 THOMSON ISI on STN
 ACCESSION NUMBER: 2000:622594 SCISEARCH
 THE GENUINE ARTICLE: 343QA
 TITLE: Vacuolar proton pyrophosphatase activity and pyrophosphate (PPi) in *Toxoplasma gondii* as possible chemotherapeutic targets
 AUTHOR: Rodrigues C O; Scott D A; Bailey B N; deSouza W; Benchimol M; Moreno B; Urbina J A; Oldfield E; Moreno S N J (Reprint)
 CORPORATE SOURCE: UNIV ILLINOIS, COLL VET MED, DEPT PATHOBIOL, MOL PARASITOL LAB, 2001 S LINCOLN AVE, URBANA, IL 61802 (Reprint); UNIV ILLINOIS, COLL VET MED, DEPT PATHOBIOL, MOL PARASITOL LAB, URBANA, IL 61802; UNIV ILLINOIS, DEPT CHEM, URBANA, IL 61801; UNIV ILLINOIS, DEPT BIOPHYS, URBANA, IL 61801; UNIV SANTA URSULA, RIO JANEIRO, BRAZIL; INST VENEZOLANO INVEST CIENT, CARACAS, VENEZUELA; UNIV FED RIO DE JANEIRO, INST BIOFIS CARLOS CHAGAS FILHO, BR-21941 RIO JANEIRO, BRAZIL
 COUNTRY OF AUTHOR: USA; BRAZIL; VENEZUELA
 SOURCE: BIOCHEMICAL JOURNAL, (1 AUG 2000) Vol. 349, Part 3, pp. 737-745.
 Publisher: PORTLAND PRESS, 59 PORTLAND PLACE, LONDON W1N 3AJ, ENGLAND.
 ISSN: 0264-6021.
 DOCUMENT TYPE: Article; Journal
 FILE SEGMENT: LIFE
 LANGUAGE: English
 REFERENCE COUNT: 50

ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS

AB The addition of PPi promoted the acidification of a subcellular compartment in cell homogenates of *Toxoplasma gondii* tachyzoites, implying the presence of a proton-translocating pyrophosphatase. The proton gradient was collapsed by addition of the K⁺/H⁺ antiporter nigericin, and was also inhibited by addition of the PP, analogue aminomethylenediphosphonate (AMDP). Both proton transport and PPi hydrolysis were dependent upon K⁺, but Na⁺ caused partial inhibition of these activities. PPi hydrolysis was sensitive in a dose-dependent manner to AMDP, imidodiphosphate, NaF and to the thiol reagent N-ethylmaleimide. This activity was unaffected by common inhibitors of phosphohydrolases, except that NaO₃V (sodium orthovanadate) **stimulated** the activity by 87%. Immunofluorescence microscopy, using antisera raised against conserved **peptide** sequences of a plant vacuolar pyrophosphatase, suggested that the pyrophosphatase in *T. gondii* tachyzoites was located in the plasma membrane and intracellular

THIS PAGE BLANK (USPTO)

vacuoles of the parasite. High-field P-31-NMR spectroscopy showed that PPI was more abundant than ATP in tachyzoites. Bisphosphonates (PPI analogues), drugs that are used in the treatment of bone diseases, inhibited proton transport and PPI hydrolysis in tachyzoite homogenates, and also inhibited intracellular proliferation of tachyzoites in tissue culture cells.

L16 ANSWER 10 OF 17 SCISEARCH COPYRIGHT 2003 THOMSON ISI on STN
 ACCESSION NUMBER: 2000:385352 SCISEARCH
 THE GENUINE ARTICLE: 314FV
 TITLE: AVP2, a sequence-divergent, K⁺-insensitive H⁺-translocating inorganic pyrophosphatase from arabidopsis
 AUTHOR: Drozdowicz Y M; Kissinger J C; Rea P A (Reprint)
 CORPORATE SOURCE: UNIV PENN, DEPT BIOL, INST PLANT SCI, PHILADELPHIA, PA 19104 (Reprint); UNIV PENN, DEPT BIOL, INST PLANT SCI, PHILADELPHIA, PA 19104
 COUNTRY OF AUTHOR: USA
 SOURCE: PLANT PHYSIOLOGY, (MAY 2000) Vol. 123, No. 1, pp. 353-362.
 Publisher: AMER SOC PLANT PHYSIOLOGISTS, 15501 MONONA DRIVE, ROCKVILLE, MD 20855.
 ISSN: 0032-0889.
 DOCUMENT TYPE: Article; Journal
 FILE SEGMENT: LIFE; AGRI
 LANGUAGE: English
 REFERENCE COUNT: 42

ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS

AB Plant vacuolar H⁺-translocating inorganic pyrophosphatases (V-PPases; EC 3.6.1.1) have been considered to constitute a family of functionally and structurally monotonous intrinsic membrane **proteins**. Typified by AVP1 (V. Sarafian, Y. Kim, R.J. Poole, P.A. Rea [1992] Proc Natl Acad Sci USA 89: 1775-1779) from Arabidopsis, all characterized plant V-PPases share greater than 84% sequence identity and catalyze K⁺-**stimulated** H⁺ translocation. Here we describe the molecular and biochemical characterization of AVP2 (accession number AF182813), a sequence-divergent (36% identical) K⁺-insensitive, Ca²⁺-hypersensitive V-PPase active in both inorganic pyrophosphate hydrolysis and H⁺ translocation. The differences between AVP2 and AVP1 provide the first indication that plant V-PPases from the same organism fall into two distinct categories. Phylogenetic analyses of these and other V-PPase sequences extend this principle by showing that AVP2, rather than being an isoform of AVP1, is but one representative of a novel category of AVP2-like (type II) V-PPases that coexist with AVP1-like (type I) V-PPases not only in plants, but also in apicomplexan protists such as the malarial parasite **Plasmodium falciparum**.

L16 ANSWER 11 OF 17 SCISEARCH COPYRIGHT 2003 THOMSON ISI on STN
 DUPLICATE 3
 ACCESSION NUMBER: 96:116360 SCISEARCH
 THE GENUINE ARTICLE: TT475
 TITLE: ISOLATION AND CHARACTERIZATION OF SEA SPONGE MYOSIN
 AUTHOR: KANZAWA N (Reprint); TAKANO OHMURO H; MARUYAMA K
 CORPORATE SOURCE: CHIBA UNIV, FAC SCI, DEPT BIOL, CHIBA 263, JAPAN (Reprint); UNIV TOKYO, FAC MED, DEPT PHARMACOL, TOKYO 113, JAPAN

THIS PAGE BLANK (USPTO)

COUNTRY OF AUTHOR: JAPAN
 SOURCE: ZOOLOGICAL SCIENCE, (DEC 1995) Vol. 12, No. 6, pp. 765-769.
 ISSN: 0289-0003.
 DOCUMENT TYPE: Article; Journal
 FILE SEGMENT: LIFE; AGRI
 LANGUAGE: ENGLISH
 REFERENCE COUNT: 17

ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS

AB Myosin was purified to a homogeneity from sea sponge, *Halichondria okadai*. The myosin consisted of 220 kDa heavy chain, 18 kDa calcium binding light chain and 21 kDa phosphorylatable light chain. Rotary shadowed images showed the two headed myosin (myosin II) with a 160 nm tail. The myosin was less soluble in a KCl solution as compared to rabbit skeletal myosin.

The K⁺-**stimulated** and Ca²⁺-**stimulated** **ATPase** activities of sea sponge myosin were 0.46 and 0.07 μ mol Pi min⁻¹ mg⁻¹, respectively. The Mg²⁺ activated myosin **ATPase** activity showed no significant enhancement by the addition of rabbit skeletal muscle actin despite that the light chain was phosphorylated by myosin light chain kinase from chicken gizzard. Sea sponge myosin 18 kDa light chain bound to Ca²⁺ ion but was not phosphorylated like *Physarum plasmodia* myosin light chains.

L16 ANSWER 12 OF 17 SCISEARCH COPYRIGHT 2003 THOMSON ISI on STN
 ACCESSION NUMBER: 95:282492 SCISEARCH
 THE GENUINE ARTICLE: QT916
 TITLE: DRUG-**STIMULATED** **ATPASE** ACTIVITY OF THE HUMAN P-GLYCOPROTEIN
 AUTHOR: SCARBOROUGH G A (Reprint)
 CORPORATE SOURCE: UNIV N CAROLINA, DEPT PHARMACOL, CHAPEL HILL, NC, 27599 (Reprint)
 COUNTRY OF AUTHOR: USA
 SOURCE: JOURNAL OF BIOENERGETICS AND BIOMEMBRANES, (FEB 1995 Vol. 27, No. 1, pp. 37-41.
) ISSN: 0145-479X.
 DOCUMENT TYPE: General Review; Journal
 FILE SEGMENT: LIFE
 LANGUAGE: ENGLISH
 REFERENCE COUNT: 41

ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS

AB The human multidrug resistance **protein**, or P-glycoprotein (Pgp), exhibits a high-capacity drug-dependent ATP hydrolytic activity that is a direct reflection of its drug transport capability. This activity is readily measured in membranes isolated from cultured insect cells infected with a baculovirus carrying the human *mdr1* cDNA. The drug-**stimulated** **ATPase** activity is a useful alternative to conventional screening systems for identifying high-affinity drug substrates of the Pgp with potential clinical value as chemosensitizers for tumor cells that have become drug resistant. Using this assay system, a variety of drugs have been directly shown to interact with the Pgp. Many of the drugs **stimulate** the Pgp **ATPase** activity, but certain drugs bind tightly to the drug-binding site of the Pgp without eliciting ATP hydrolysis. Either class of drugs may be useful as chemosensitizing agents. The baculovirus/insect cell Pgp **ATPase** assay system may also facilitate future studies

THIS PAGE BLANK (USPTO)

of the molecular structure and mechanism of the Pgp.

L16 ANSWER 13 OF 17 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC. on STN

ACCESSION NUMBER: 1993:346268 BIOSIS
DOCUMENT NUMBER: PREV199396043268
TITLE: Absence of transitory (Ca-2+)-i flux during early in vitro metacyclogenesis of Trypanosoma cruzi.
AUTHOR(S): Krassner, Stuart M. (1); Chang, Johnny; Pak, Sung; Luc, Kim-Oanh; Granger, Barbara
CORPORATE SOURCE: (1) Dep. Dev. Cell Biol., Univ. Calif., Irvine, CA 92717 USA
SOURCE: Journal of Eukaryotic Microbiology, (1993) Vol. 40, No. 2, pp. 224-230.
ISSN: 1066-5234.
DOCUMENT TYPE: Article
LANGUAGE: English

AB The phorbol ester TPA (phorbol 12-myristate 13-acetate) substitutes for CO-2 as an agonist for transforming Trypanosoma cruzi epimastigotes to the metacyclic trypomastigote stage in a starvation medium consisting of phosphate buffered saline + 10 mM proline, 10 mM sodium acetate and 0.035% NaHCO-3. Since TPA is thought to **stimulate protein kinase C** by mimicking the activity of the secondary messenger diacylglycerol, the above result suggested that T. cruzi metacyclogenesis could be activated by a Ca-2+-dependent **protein kinase C** signal induction pathway. Accordingly, cytosolic calcium Bux ((Ca-2+)-i) in epimastigotes, activated with 5% CO-2 or TPA (10⁻⁷ M), was measured with the Ca-2+ molecular probe, fluo-3AM. In addition, (Ca-2+)-i was measured in cells incubated with putative metacyclogenetic agonists (e.g. proline, glutamate, bioamines, ionophores and catecholamines). None of the compounds studied, except for EGTA, affected cytosolic Ca-2+ levels. Control assays with 11 mu-M thapsigargin, which mobilizes noncytoplasmic Ca-2+ stores by inhibiting endoplasmic reticulum Ca-2+-ATPase, validated our fluorometric assay procedure. Although thapsigargin significantly increases cytoplasmic Ca-2+ fluorescence, it has no effect on transformation. The **protein kinase C** inhibitors staurosporine, H-7 and HA 1004 were tested for their effect on T. cruzi metacyclogenesis. Low concentrations of staurosporine and HA 1004 significantly elevated Peru strain transformation while H-7 had no effect on Peru strain metacyclogenesis. Inhibitor H-7 did significantly depress CL transformation. The results indicate that induction of T. cruzi metacyclic trypomastigote formation by CO-2 and TPA is not accompanied by changes in cytosolic Ca-2+ and do not provide supporting evidence for participation of a **protein kinase C**-mediated phosphoinositide cascade in metacyclogenesis.

L16 ANSWER 14 OF 17 MEDLINE on STN DUPLICATE 4
ACCESSION NUMBER: 92042085 MEDLINE
DOCUMENT NUMBER: 92042085 PubMed ID: 1834662
TITLE: **Stimulation** of the interaction between actin and myosin by Physarum caldesmon-like **protein** and smooth muscle caldesmon.
AUTHOR: Ishikawa R; Okagaki T; Higashi-Fujime S; Kohama K
CORPORATE SOURCE: Department of Pharmacology, Gunma University School of Medicine, Japan.
SOURCE: JOURNAL OF BIOLOGICAL CHEMISTRY, (1991 Nov 15) 266

THIS PAGE BLANK (USPTO)

10/006780

(32) 21784-90.
Journal code: 2985121R. ISSN: 0021-9258.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199112
ENTRY DATE: Entered STN: 19920124
Last Updated on STN: 19970203
Entered Medline: 19911223

AB We have purified an actin-binding **protein** from the **plasmodia** of a lower eukaryote, *Physarum polycephalum*, with an apparent molecular mass of 210,000 daltons on sodium dodecyl sulfate-polyacrylamide gel electrophoresis. This **protein** bound to actin filaments with a stoichiometry of 1:7-8 in a $\text{Ca}(2+)$ -calmodulin-dependent manner. Antibody raised against caldesmon from smooth muscle cross-reacted with the 210-kDa **protein**. In vitro motility assay revealed that the 210-kDa **protein** increased the sliding velocity of actin filaments on *Physarum* myosin. The 210-kDa **protein** more than doubled the actin-activated **ATPase** activity of *Physarum* myosin under comparative conditions of in vitro motility assay. Further increases in the concentration of the 210-kDa **protein** decreased its **stimulatory** effects. $\text{Ca}(2+)$ -calmodulin prevented the **stimulatory** effects of the 210-kDa **protein**. Unexpectedly, smooth muscle caldesmon also increased the sliding velocity of actin filaments on smooth muscle myosin at lower concentrations. The well-known inhibitory effect of smooth muscle caldesmon on the actin-myosin interaction was observed with this motility assay when the concentration of the caldesmon was increased further. The **stimulatory** and inhibitory effects were confirmed by measurements of actin-activated **ATPase** activity of smooth muscle myosin. From estimations of the intracellular concentrations of the 210-kDa **protein** and smooth muscle caldesmon in vivo, it appears that effects of the former and the latter on actin-myosin interactions in vivo are **stimulatory** and inhibitory, respectively.

L16 ANSWER 15 OF 17 SCISEARCH COPYRIGHT 2003 THOMSON ISI on STN
ACCESSION NUMBER: 91:631866 SCISEARCH
THE GENUINE ARTICLE: GP804
TITLE: **STIMULATION OF THE INTERACTION BETWEEN
ACTIN AND MYOSIN BY PHYSARUM CALDESMON-LIKE
PROTEIN AND SMOOTH-MUSCLE CALDESMON**
AUTHOR: ISHIKAWA R (Reprint); OKAGAKI T; HIGASHIFUJIME S;
KOHAMA K
CORPORATE SOURCE: GUNMA UNIV, SCH MED, DEPT PHARMACOL, MAEBASHI, GUNMA
371, JAPAN (Reprint); NIHON UNIV, PHYS SCI LABS,
FUNABASHI, CHIBA 274, JAPAN; NAGOYA UNIV, FAC SCI,
DEPT MOLEC BIOL, NAGOYA, AICHI 464, JAPAN
COUNTRY OF AUTHOR: JAPAN
SOURCE: JOURNAL OF BIOLOGICAL CHEMISTRY, (1991) Vol. 266,
No. 32, pp. 21784-21790.
DOCUMENT TYPE: Article; Journal
FILE SEGMENT: LIFE
LANGUAGE: ENGLISH
REFERENCE COUNT: 30
ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS

THIS PAGE BLANK (USPTO)

AB We have purified an actin-binding **protein** from the **plasmodia** of a lower eukaryote, *Physarum polycephalum*, with an apparent molecular mass of 210,000 daltons on sodium dodecyl sulfate-polyacrylamide gel electrophoresis. This **protein** bound to actin filaments with a stoichiometry of 1:7-8 in a Ca^{2+} -calmodulin-dependent manner. Antibody raised against caldesmon from smooth muscle cross-reacted with the 210-kDa **protein**. In vitro motility assay revealed that the 210-kDa **protein** increased the sliding velocity of actin filaments on *Physarum* myosin. The 210-kDa **protein** more than doubled the actin-activated **ATPase** activity of *Physarum* myosin under comparative conditions of in vitro motility assay. Further increases in the concentration of the 210-kDa **protein** decreased its **stimulatory** effects. Ca^{2+} -calmodulin prevented the **stimulatory** effects of the 210-kDa **protein**. Unexpectedly, smooth muscle caldesmon also increased the sliding velocity of actin filaments on smooth muscle myosin at lower concentrations. The well-known inhibitory effect of smooth muscle caldesmon on the actin-myosin interaction was observed with this motility assay when the concentration of the caldesmon was increased further. The **stimulatory** and inhibitory effects were confirmed by measurements of actin-activated **ATPase** activity of smooth muscle myosin. From estimations of the intracellular concentrations of the 210-kDa **protein** and smooth muscle caldesmon in vivo, it appears that effects of the former and the latter on actin-myosin interactions in vivo are **stimulatory** and inhibitory, respectively.

L16 ANSWER 16 OF 17 MEDLINE on STN DUPLICATE 5
 ACCESSION NUMBER: 89040035 MEDLINE
 DOCUMENT NUMBER: 89040035 PubMed ID: 2972931
 TITLE: Purification of **Plasmodium falciparum** digestive vacuoles and partial characterization of the vacuolar membrane **ATPase**.
 AUTHOR: Choi I; Mego J L
 CORPORATE SOURCE: Department of Biology, University of Alabama, Tuscaloosa 35487.
 SOURCE: MOLECULAR AND BIOCHEMICAL PARASITOLOGY, (1988 Oct) 31 (1) 71-8.
 Journal code: 8006324. ISSN: 0166-6851.
 PUB. COUNTRY: Netherlands
 DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
 LANGUAGE: English
 FILE SEGMENT: Priority Journals
 ENTRY MONTH: 198812
 ENTRY DATE: Entered STN: 19900308
 Last Updated on STN: 19900308
 Entered Medline: 19881220

AB **Plasmodium falciparum** digestive vacuoles containing ferric oxide granules were purified from parasite homogenates by centrifugation on discontinuous sucrose gradients. Digestive vacuole membranes prepared by osmotic lysis and washed with KCl showed no detectable contamination by erythrocyte membrane **proteins** and only minimal contamination by non-vacuolar parasite **proteins**. Purified vacuolar membranes were 2.6-fold enriched in total parasite membrane **ATPase** activity. This **ATPase** was optimally active at pH 7 in the

THIS PAGE BLANK (USPTO)

presence of at least 2 mM Mg²⁺. Ca²⁺ and Mn²⁺ were approximately 80-90% as effective as Mg²⁺, and Zn²⁺, Co²⁺ and Fe²⁺ also exerted some **stimulatory** effect. The vacuolar membrane also hydrolyzed GTP, UTP, CTP and ADP, but AMP and 3',5'-cyclic AMP were hydrolyzed only one-tenth as effectively as ATP. The **ATPase** was unaffected by vanadate, ouabain or oligomycin but was significantly inhibited by the proton pump inhibitors NEM and NBD-Cl. Of 6 antimalarial drugs tested, quinine and quinacrine were the most effective inhibitors and mefloquine was the least effective.

L16 ANSWER 17 OF 17 EMBASE COPYRIGHT 2003 ELSEVIER INC. ALL RIGHTS RESERVED. on STN

ACCESSION NUMBER: 74138989 EMBASE

DOCUMENT NUMBER: 1974138989

TITLE: The biosynthesis of plasmodial myosin during starvation of *Physarum polycephalum*.

AUTHOR: White F.H.; Lascelles J.

CORPORATE SOURCE: Dept. Bacteriol., Univ. California, Los Angeles, Calif. 90025, United States

SOURCE: Biochemical Journal, (1973) 135/4 (639-647).
CODEN: BIJOAK

DOCUMENT TYPE: Journal

FILE SEGMENT: 029 Clinical Biochemistry

LANGUAGE: English

AB The actomyosin **protein** complex of *Physarum polycephalum* was prepared from vegetative and starved **plasmodia**. The yield of actomyosin per unit wet wt was the same from both types of **plasmodia**. Myosin was resolved from the complex by gel filtration and purified by ion exchange chromatography. The Ca²⁺ **stimulated adenosine triphosphatase** activities of myosin preparations from vegetative and starved **plasmodia** were not appreciably different. Synthesis of myosin de novo was shown to occur during the starvation phase of the life cycle by the isolation of labelled myosin preparations from **plasmodia** starved in the presence of [2-¹⁴C]glycine. Fractionation of polyacrylamide gels after gel filtration of labelled myosin confirmed the presence of label in the **adenosine triphosphatase** active myosin band. It is concluded that during starvation myosin synthesis continues although there is a net loss of approx. 50% of the total **protein**. Sodium dodecyl sulphate polyacrylamide gel electrophoresis of *Physarum* myosin showed the presence of low molecular weight components of the molecule, similar to those of muscle myosins. The content and composition of the free amino acid pool to *Physarum* was measured at various time intervals during the vegetative and starvation phases of the life cycle.

(FILE 'MEDLINE' ENTERED AT 14:43:41 ON 07 OCT 2003)
L29 12195 SEA FILE=MEDLINE ABB=ON PLU=ON "PLASMODIUM FALCIPARUM"/
CT
L30 28410 SEA FILE=MEDLINE ABB=ON PLU=ON ADENOSINETRIPHOSPHATASE/
CT
L31 17 SEA FILE=MEDLINE ABB=ON PLU=ON L29 AND L30

L31 ANSWER 1 OF 17 MEDLINE on STN

AN 2002675294 MEDLINE

TI Transport processes in *Plasmodium falciparum*-infected erythrocytes:

THIS PAGE BLANK (USPTO)

- potential as new drug targets.
- AU Krishna Sanjeev; Eckstein-Ludwig Ursula; Joet Thierry; Uhlemann Anne-Catrin; Morin Christophe; Webb Richard; Woodrow Charles; Kun Jorgen F J; Kremsner Peter G
- SO INTERNATIONAL JOURNAL FOR PARASITOLOGY, (2002 Dec 4) 32 (13) 1567-73. Ref: 39
Journal code: 0314024. ISSN: 0020-7519.
- AB Plasmodium falciparum infection induces alterations in the transport properties of infected erythrocytes that have recently been defined using electrophysiological techniques. Mechanisms responsible for transport of substrates into intraerythrocytic parasites have also been clarified by studies of three substrate-specific (hexose, nucleoside and aquaglyceroporin) parasite plasma membrane transporters. These have been characterised functionally using the *Xenopus laevis* oocyte heterologous expression system. The same expression system is currently being used to define the function of parasite 'P' type ATPases responsible for intraparasitic [Ca(2+)] homeostasis. We review studies on these transport processes and examine their potential as novel drug targets.
- L31 ANSWER 2 OF 17 MEDLINE on STN
- AN 2001479314 MEDLINE
- TI Characterization of P-type ATPase 3 in *Plasmodium falciparum*.
- AU Rozmajzl P J; Kimura M; Woodrow C J; Krishna S; Meade J C
- SO MOLECULAR AND BIOCHEMICAL PARASITOLOGY, (2001 Sep 3) 116 (2) 117-26. Journal code: 8006324. ISSN: 0166-6851.
- AB We report the nucleotide sequence, derived amino acid sequence and expression profile of P-type ATPase 3 (PfATPase3) from *Plasmodium falciparum*. An open reading frame of 7362 nucleotides, interrupted by a single intron of 168 nt, encoded a protein product of 2394 amino acids with a predicted MW of 282791 Da. Hydropathy analysis of PfATPase3 revealed six amino-terminal and six carboxyl-terminal membrane spanning regions (M1-12) flanking a large hydrophilic domain with a smaller hydrophilic loop between M4 and M5. Based on a phylogenetic comparison of conserved domains present in P-type ATPases from other organisms, PfATPase3 resembled a Type-V ATPase for which the transport affinity is unknown. The PfATPase3 topology was interrupted by four regions, termed 'inserts', unique to malarial P-type ATPases, which were high in asparagine residues and charged amino acids (inserts I1-I4). Inserts I1 and I3 also contained repeated amino acid motifs. The number and composition of repeated amino acid motifs in insert I3 were variable in seven *P. falciparum* strains tested. PfATPase3 was 80.2% similar to the non-insert portions of *P. yoelii* ATPase3, although their inserts differed in length and composition. PfATPase3 mRNA was most abundant relative to beta-tubulin during the latter half of the erythrocytic cycle and was also present in gametocytes. Using affinity-purified antibody to a 14 amino acid PfATPase3 epitope, a 260 kDa protein was detected by Western analysis. Based on immunofluorescence, the PfATPase3 protein was located intracellularly in gametocytes and, to a lesser extent, in late erythrocytic stages.
- L31 ANSWER 3 OF 17 MEDLINE on STN
- AN 2000386953 MEDLINE
- TI Molecular cloning of a gene encoding a 20S proteasome beta subunit from *Plasmodium falciparum*.
- AU Li G D; Li J L; Mugthi M; Ward S A

THIS PAGE BLANK (USPTO)

- SO INTERNATIONAL JOURNAL FOR PARASITOLOGY, (2000 May) 30 (6) 729-33.
Journal code: 0314024. ISSN: 0020-7519.
- AB A novel gene was cloned from *Plasmodium falciparum*. Database searches indicated this gene to be a member of the 20S proteasome beta-subunit family. Comparison of the gene's genomic DNA sequence with cDNA sequence revealed a 156-bp intron 85 bp downstream from the start codon. The nucleotide sequence of the gene contains one open reading frame encoding 265 amino acids with a predicted molecular mass of 30.9 kDa and a pI of 6.2. Northern blot analysis showed the transcript size to be approximately 1.6 kb indicating that some 800 bp of the transcript is non-coding.
- L31 ANSWER 4 OF 17 MEDLINE on STN
- AN 2000002518 MEDLINE
- TI Cloning and partial characterization of the proteasome S4 ATPase from *Plasmodium falciparum*.
- AU Certad G; Abraham A; Georges E
- SO EXPERIMENTAL PARASITOLOGY, (1999 Nov) 93 (3) 123-31.
Journal code: 0370713. ISSN: 0014-4894.
- AB Certad, G., Abraham, A., and Georges, E. 1999. Cloning and Partial characterization of the proteasome S4 ATPase from *Plasmodium falciparum*. *Experimental Parasitology* 93, 123-131. The ATP-ubiquitin-proteasome pathway mediates the nonlysosomal degradation of cytosolic proteins in eukaryotic cells. The activities of this pathway have been shown to regulate cell growth and differentiation through modulation of regulatory proteins. The proteasome is a large complex consisting of two multisubunit structures, the 20S and 19S(PA700) or P28 complexes, that combine to form the 26S particles. In this study, we describe the cloning of a cDNA encoding the proteasome subunit 4 ATPase homologue from *Plasmodium falciparum* (PFS4). Analysis of the PFS4 cDNA sequence shows an open reading frame encoding a deduced protein of 455 amino acids. Moreover, comparison of PFS4 cDNA sequence to that of genomic fragments encoding PFS4 showed identical sequences with no detectable introns. Database searches revealed a high sequence identity to those of rice, yeast, mouse, *Drosophila*, and human S4 ATPases. However, PFS4 contains two unique inserts of nine and seven amino acid residues in the N-terminal domain. Interestingly, only the rice S4 contains the latter (seven amino acids) insert with four identical amino acids. In vitro expression of the full-length cDNA encoding the PFS4, using a transcription-translation-coupled reticulocyte lysate, shows a 50-kDa [(35)S]methionine-labeled protein which was immunoprecipitated with PFS4 anti-peptide antiserum. Southern blot analysis of genomic DNA digests shows a single gene copy of PFS4 in *P. falciparum*. Of interest was the effect of the proteasome-specific natural product, lactacystin, on the growth of the parasite, with IC(50) values of 0.6-0.92 microM. The latter IC(50) values of lactacystin for different clones of *P. falciparum* are comparable to those obtained for mammalian cell lines (0.65 microM), suggesting the presence of a conserved proteasome complex. Moreover, lactacystin was equally toxic to drug-sensitive and resistant parasites.
Copyright 1999 Academic Press.
- L31 ANSWER 5 OF 17 MEDLINE on STN
- AN 97418800 MEDLINE
- TI A *Plasmodium falciparum* homologue of the ATPase subunit of a multi-protein complex involved in chromatin remodelling for

THIS PAGE BLANK (USPTO)

- transcription.
- AU Ji D D; Arnot D E
- SO MOLECULAR AND BIOCHEMICAL PARASITOLOGY, (1997 Sep) 88 (1-2) 151-62.
Journal code: 8006324. ISSN: 0166-6851.
- AB A Plasmodium falciparum homologue of one of the components of a chromatin-remodelling complex which controls binding of transcription factors to nucleosome core particles has been cloned and characterised. The gene encodes 1422 amino acids with an estimated molecular mass of 167 kDa. The protein, SNF2L, shares 60% amino acid identity in its conserved DNA-dependent ATPase domain with yeast transcription factors originally identified by characterising mating type switch mutants. It also contains sequences related to the so-called SWI3, ADA2, N-CoR and TFIIIB B" or SANT DNA binding domains which are characteristic of these transcriptional activation factors. The SNF2L gene has two short introns in the 3' region of the coding sequence of the gene and is transcribed into a single approximately 6.5 kb messenger RNA species which is present throughout the asexual stages of the cell cycle. Southern blotting and pulsed field gel electrophoresis experiments show that SNF2L is a single copy gene. located on P. falciparum chromosome 11.
- L31 ANSWER 6 OF 17 MEDLINE on STN
- AN 96408665 MEDLINE
- TI Analysis of a cation-transporting ATPase of Plasmodium falciparum.
- AU Dyer M; Jackson M; McWhinney C; Zhao G; Mikkelsen R
- SO MOLECULAR AND BIOCHEMICAL PARASITOLOGY, (1996 Jun) 78 (1-2) 1-12.
Journal code: 8006324. ISSN: 0166-6851.
- AB We have cloned and characterised one gene, PfATPase4 which encodes a P-type ATPase containing all the primary sequence motifs characteristic of this class of transmembrane ion transporters, and also a fragment of a second P. falciparum P-type ATPase pseudogene (PfATPase5). Analysis of conserved domains and motifs of specific ATPases reveals that PfATPase4 is most analogous to Ca²⁺ ATPases of the endoplasmic reticulum. The PfATPase4 gene gives rise to a transcript of 8 kb shortly after erythrocyte invasion. Although this mRNA is not detected in later stages, the protein detected immunologically at 190 kDa persists throughout and is detected in free merozoites. Immunofluorescence microscopy reveals that the PfATPase4 protein is concentrated in discrete compartments at the periphery of the parasite. Detailed sequence and structural analyses of these and the other P-type ATPases of P. falciparum described previously, reveals that they comprise an unusual family in several respects. Firstly, the large number of non-homologous genes so far characterised reflects the complexities of ionic regulation in the diverse environments encountered by the parasite. Secondly, the plasmodial P-type ATPase family may be classified both at primary sequence and structural levels into two distinct groups-those typical of P-type ATPases (including PfATPase4) and those which are much more divergent. A third complexity is illustrated by the fact that one of the other members [1] here termed PfATPase6, has an even greater similarity to the sarcoplasmic reticulum Ca²⁺ ATPases than does PfATPase4, which raises questions about the possible functional relationship between these two members.
- L31 ANSWER 7 OF 17 MEDLINE on STN
- AN 95309711 MEDLINE

THIS PAGE BLANK (USPTO)

- TI Cloning of a new cation ATPase from *Plasmodium falciparum*: conservation of critical amino acids involved in calcium binding in mammalian organellar Ca(2+)-ATPases.
- AU Trottein F; Thompson J; Cowman A F
- SO GENE, (1995 May 26) 158 (1) 133-7.
Journal code: 7706761. ISSN: 0378-1119.
- AB In order to study molecules that may be involved in pH gradient formation in *Plasmodium*, we have identified a novel cation-translocating ATPase (P-type ATPase) gene from *P. falciparum* (Pf). We report the full-length nucleotide and deduced amino acid (aa) sequences of this gene that we called PfATPase4. The PfATPase4 protein shares features with the different members of eukaryotic P-type ATPases, such as a similar transmembrane (TM) organization and aa identity in functionally important regions. Interestingly, the PfATPase4 protein possesses conserved aa involved in calcium binding in mammalian organellar Ca(2+)-ATPases.
- L31 ANSWER 8 OF 17 MEDLINE on STN
- AN 95154293 MEDLINE
- TI Molecular cloning and sequence of two novel P-type adenosinetriphosphatases from *Plasmodium falciparum*.
- AU Trottein F; Cowman A F
- SO EUROPEAN JOURNAL OF BIOCHEMISTRY, (1995 Jan 15) 227 (1-2) 214-25.
Journal code: 0107600. ISSN: 0014-2956.
- AB We have identified two novel P-type ATPase genes from *Plasmodium falciparum* and report the full-length nucleotide and derived amino acid sequence of the ATPase2 gene from *P. falciparum* (PfATPase2). PfATPase2 is phylogenetically remote from the different members of prokaryotic and mammalian P-type ATPases but shares features with a putative membrane-spanning Ca2+ ATPase involved in ribosome function in yeast. PfATPase2 is expressed during the intraerythrocytic life cycle of the parasite and appears to be required in the late stages of its asexual development. We also present the partial sequence of another malarial gene displaying sequence similarity with the family of P-type transporting ATPases (PfATPase4). We have analysed the organisation of the genes encoding the P-type ATPases of *P. falciparum* and show that they are a highly dispersed gene family.
- L31 ANSWER 9 OF 17 MEDLINE on STN
- AN 95021518 MEDLINE
- TI Cloning and characterization of the vacuolar ATPase B subunit from *Plasmodium falciparum*.
- AU Karcz S R; Herrmann V R; Trottein F; Cowman A F
- SO MOLECULAR AND BIOCHEMICAL PARASITOLOGY, (1994 May) 65 (1) 123-33.
Journal code: 8006324. ISSN: 0166-6851.
- AB The transvacuolar pH gradient determines, to a significant extent, the distribution of the antimalarial drug chloroquine in *Plasmodium falciparum*. A proton pump, similar to the vacuolar ATPase found in many cell types, appears to regulate a pH gradient across the membranes of acidic compartments of the parasite. In order to understand and define the components involved in the maintenance of the vacuolar pH gradient, we have cloned and characterized a gene, designated VAP B, encoding a *P. falciparum* homologue of the B subunit of the vacuolar ATPase. The VAP B gene encodes a protein of 494 amino acids which has between 69% and 74% amino acid identity with the sequences of vacuolar ATPase B subunits of other organisms. The VAP B gene exists as a single copy gene on chromosome 4 that gives rise to a RNA transcript of 2.4 kb. Antibodies raised to the

THIS PAGE BLANK (USPTO)

VAP B protein react specifically with a protein of 56-kDa, consistent with the size predicted from the gene sequence and with the homologous protein from other organisms. The 56-kDa protein is expressed throughout the asexual life cycle and subcellular localization by indirect immunofluorescence shows that the protein has a heterogeneous distribution over most of the parasite. This suggests that the function of the vacuolar proton ATPase is not confined to the regulation of the pH of the digestive vacuole.

- L31 ANSWER 10 OF 17 MEDLINE on STN
 AN 94131050 MEDLINE
 TI Plasmodium falciparum: further characterization of putative cation ATPases.
 AU Krishna S; Cowan G M; Robson K J; Meade J C
 SO EXPERIMENTAL PARASITOLOGY, (1994 Feb) 78 (1) 113-7.
 Journal code: 0370713. ISSN: 0014-4894.
- L31 ANSWER 11 OF 17 MEDLINE on STN
 AN 93241225 MEDLINE
 TI Cloning and characterization of a vacuolar ATPase A subunit homologue from Plasmodium falciparum.
 AU Karcz S R; Herrmann V R; Cowman A F
 SO MOLECULAR AND BIOCHEMICAL PARASITOLOGY, (1993 Apr) 58 (2) 333-44.
 Journal code: 8006324. ISSN: 0166-6851.
- AB The distribution of the antimalarial drug chloroquine is determined to a significant extent by a transvacuolar pH gradient in Plasmodium falciparum. A proton pump similar to the vacuolar ATPase found in many cell types has been suggested to maintain a pH gradient across the membranes of acidic compartments in the parasite. In order to understand and define the components involved in the mechanism of acidification of parasite vesicles, we have cloned and characterized a gene, designated VAP-A, encoding a P. falciparum homologue of the catalytic A subunit of the vacuolar ATPase. The VAP-A gene encodes a polypeptide of 611 amino acids which shows between 56 to 61% amino acid identity over its entire length with the sequences of vacuolar ATPase A subunits from several species. The VAP-A gene exists as a single copy gene on P. falciparum chromosome 13 and gives rise to a transcript of 3.7 kb. Antibodies raised against a VAP-A gene segment expressed in Escherichia coli react specifically with a 67-kDa polypeptide, consistent with the size predicted from the sequence and with the size of the corresponding polypeptide in other organisms. The 67-kDa protein is present throughout the asexual erythrocytic cycle and is expressed at similar levels in 5 P. falciparum isolates of differing chloroquine sensitivity. Sequence analysis of the coding region of the VAP-A gene from 2 chloroquine-sensitive and 3 chloroquine-resistant isolates has shown no changes that are linked to chloroquine resistance. Therefore, a proposed chloroquine resistance-linked vacuolar acidification defect does not involve mutations in the VAP-A gene in the isolates we have studied.
- L31 ANSWER 12 OF 17 MEDLINE on STN
 AN 93132070 MEDLINE
 TI A family of cation ATPase-like molecules from Plasmodium falciparum.
 AU Krishna S; Cowan G; Meade J C; Wells R A; Stringer J R; Robson K J
 SO JOURNAL OF CELL BIOLOGY, (1993 Jan) 120 (2) 385-98.
 Journal code: 0375356. ISSN: 0021-9525.
- AB We report the nucleotide and derived amino acid sequence of the

THIS PAGE BLANK (USPTO)

ATPase 1 gene from *Plasmodium falciparum*. The amino acid sequence shares homology with the family of "P"-type cation translocating ATPases in conserved regions important for nucleotide binding, conformational change, or phosphorylation. The gene, which is present on chromosome 5, has a product longer than any other reported for a P-type ATPase. Interstrain analysis from 12 parasite isolates by the polymerase chain reaction reveals that a 330-bp nucleotide sequence encoding three cytoplasmic regions conserved in cation ATPases (regions a-c) is of constant length. By contrast, another 360-bp sequence which is one of four regions we refer to as "inserts" contains arrays of tandem repeats which show length variation between different parasite isolates. Polymorphism results from differences in the number and types of repeat motif contained in this insert. Inserts are divergent in sequence from other P-type ATPases and share features in common with many malarial antigens. Studies using RNA from the erythrocytic stages of the malarial life cycle suggest that ATPase 1 (including the sequence which encodes tandem repeats) is expressed at the large ring stage of development. Immunolocalization has identified ATPase 1 to be in the region of the parasite plasma membrane and pigment body. These findings suggest a possible model for the genesis of malarial antigens.

- L31 ANSWER 13 OF 17 MEDLINE on STN
 AN 92134363 MEDLINE
 TI Energy dependence of chloroquine accumulation and chloroquine efflux in *Plasmodium falciparum*.
 AU Krogstad D J; Gluzman I Y; Herwaldt B L; Schlesinger P H; Wellems T E
 SO BIOCHEMICAL PHARMACOLOGY, (1992 Jan 9) 43 (1) 57-62.
 Journal code: 0101032. ISSN: 0006-2952.
 AB Chloroquine inhibits the growth of susceptible malaria parasites at low (nanomolar) concentrations because of an energy-requiring drug-concentrating mechanism in the parasite secondary lysosome (food vacuole) which is dependent on the acidification of that vesicle. Chloroquine resistance results from another energy-requiring process: efflux of chloroquine from the resistant parasite with a half-time of 2 min. Chloroquine efflux is inhibited reversibly by the removal of metabolizable substrate (glucose); it is also reduced by the ATPase inhibitor vanadate. These results suggest that chloroquine efflux is an energy-requiring process dependent on the generation and hydrolysis of ATP. Chloroquine efflux cannot be explained by differences in drug accumulation between chloroquine-susceptible and -resistant parasites because the 40-50-fold difference in initial efflux rates between -susceptible and -resistant parasites is unchanged when both parasites contain the same amount of chloroquine. Although chloroquine efflux is phenotypically similar to the efflux of anticancer drugs from multidrug-resistant (mdr) mammalian cells, it is not linked to either of the mdr-like genes of the parasite.
- L31 ANSWER 14 OF 17 MEDLINE on STN
 AN 91101662 MEDLINE
 TI Accumulation of chloroquine by membrane preparations from *Plasmodium falciparum*.
 AU Herwaldt B L; Schlesinger P H; Krogstad D J
 SO MOLECULAR AND BIOCHEMICAL PARASITOLOGY, (1990 Sep-Oct) 42 (2) 257-67.
 Journal code: 8006324. ISSN: 0166-6851.

THIS PAGE BLANK (USPTO)

AB Chloroquine susceptibility and resistance have been associated respectively with the uptake and efflux of chloroquine by *Plasmodium falciparum*. We made membrane preparations from parasitized and unparasitized red cells in order to study chloroquine accumulation in a cell-free system. The accumulation of [³H]chloroquine by these preparations is inhibited by unlabeled chloroquine and thus is specific. Only membranes from parasitized red cells demonstrate time-dependent chloroquine accumulation; membranes from unparasitized red cells do not. Chloroquine accumulation is eliminated by detergent (0.05% Triton X-100) and reduced by a hypertonic medium, consistent with accumulation inside membrane vesicles rather than binding to membranes. Accumulation is energy dependent; it has a specific requirement for ATP, which cannot be replaced with GTP, CTP, UTP, TTP or ADP, an apparent *K_m* of 21 μ M and an apparent *V_{max}* of 4.6 pmol (mg protein)⁻¹ h⁻¹. Vesicle acidification is MgATP dependent, and is reversed by NH₄Cl. Chloroquine accumulation is inhibited by reduced medium pH, N-ethylmaleimide or oligomycin, but not by vanadate or ouabain. These studies demonstrate that membrane vesicles prepared from parasitized red cells provide a model system for the study of chloroquine accumulation by *P. falciparum*.

L31 ANSWER 15 OF 17 MEDLINE on STN

AN 91077466 MEDLINE

TI Ion metabolism in malaria-infected erythrocytes.

AU Tanabe K

SO BLOOD CELLS, (1990) 16 (2-3) 437-49. Ref: 61
Journal code: 7513567. ISSN: 0340-4684.

AB Malaria parasites of the genus *Plasmodium* spend much of their asexual life cycle inside the erythrocytes of their vertebrate hosts. Parasites presumably have to exploit metabolic and transport mechanisms to adapt themselves to the host erythrocyte's physicochemical environment. This review surveys the metabolism and transport of Ca²⁺, alkali cations, and H⁺ in malaria-infected erythrocytes. The Ca²⁺ content of *Plasmodium*-infected erythrocytes increases as the parasite matures. An increase in the influx of extracellular Ca²⁺ into infected erythrocytes is evident at later stages of parasite development. In infected erythrocytes, Ca²⁺ is almost exclusively localized in the parasite compartment and changes but little in the cytosol of the host cell. The importance of Ca²⁺ in supporting the growth of intraerythrocytic parasites and the invasion of erythrocytes by the merozoite has been assessed by depletion of extracellular Ca²⁺ with chelators, or by disturbance of the metabolism and transport of Ca²⁺ with a variety of Ca²⁺ modulators. Membranes of malaria-infected erythrocytes change their permeability to alkali cations. Hence, levels of K⁺ decrease and levels of Na⁺ increase in the cytosol of infected erythrocytes. Intraerythrocytic parasites maintain a high K⁺, low Na⁺ state, suggesting a mechanism for transporting K⁺ inward and Na⁺ outward against concentration gradients of the alkali cations across the parasite plasma membrane and/or the parasitophorous vacuole membrane (PVM). Concomitantly, *P. falciparum* can grow in Na⁺-enriched human erythrocytes. Experimental evidence suggests that *Plasmodium* possesses in its plasma membrane a proton pump which is very sensitive to orthovanadate, carbonylcyanide *m*-chlorophenylhydrazone, a protonophore, and dicyclohexylcarbodiimide, an inhibitor of H⁺-ATPase, but is only slightly sensitive to inhibitors of bacterial and mitochondrial respiration, such as antimycin A, CN⁻,

THIS PAGE BLANK (USPTO)

or N3-, and ouabain, a Na⁺, K⁽⁺⁾-ATPase inhibitor. By operating this proton pump, parasites extrude H⁺ and thus generate an electrochemical gradient of protons (an internal negative membrane potential and a concentration gradient of protons) across the parasite plasma membrane. The electrochemical gradient apparently drives inward movement of Ca²⁺ and, possibly, glucose from the cytosol of infected erythrocytes. Little is known about the transport properties of the PVM. Recent sequence studies suggest that Plasmodium contains a cation-transporting ATPase which exhibits a high homology to the Ca²⁺(+)-ATPase of rabbit skeletal muscle sarcoplasmic reticulum. (ABSTRACT TRUNCATED AT 250 WORDS)

L31 ANSWER 16 OF 17 MEDLINE on STN
 AN 89040035 MEDLINE
 TI Purification of Plasmodium falciparum digestive vacuoles and partial characterization of the vacuolar membrane ATPase.
 AU Choi I; Mego J L
 SO MOLECULAR AND BIOCHEMICAL PARASITOLOGY, (1988 Oct) 31 (1) 71-8.
 Journal code: 8006324. ISSN: 0166-6851.
 AB Plasmodium falciparum digestive vacuoles containing ferric oxide granules were purified from parasite homogenates by centrifugation on discontinuous sucrose gradients. Digestive vacuole membranes prepared by osmotic lysis and washed with KCl showed no detectable contamination by erythrocyte membrane proteins and only minimal contamination by non-vacuolar parasite proteins. Purified vacuolar membranes were 2.6-fold enriched in total parasite membrane ATPase activity. This ATPase was optimally active at pH 7 in the presence of at least 2 mM Mg²⁺. Ca²⁺ and Mn²⁺ were approximately 80-90% as effective as Mg²⁺, and Zn²⁺, Co²⁺ and Fe²⁺ also exerted some stimulatory effect. The vacuolar membrane also hydrolyzed GTP, UTP, CTP and ADP, but AMP and 3',5'-cyclic AMP were hydrolyzed only one-tenth as effectively as ATP. The ATPase was unaffected by vanadate, ouabain or oligomycin but was significantly inhibited by the proton pump inhibitors NEM and NBD-Cl. Of 6 antimalarial drugs tested, quinine and quinacrine were the most effective inhibitors and mefloquine was the least effective.

L31 ANSWER 17 OF 17 MEDLINE on STN
 AN 74085067 MEDLINE
 TI Bongkrekic acid and the adenosinetriphosphate requirement of malaria parasites.
 AU Trager W
 SO EXPERIMENTAL PARASITOLOGY, (1973 Dec) 34 (3) 412-6.
 Journal code: 0370713. ISSN: 0014-4894.

FILE 'HCAPLUS' ENTERED AT 14:45:35 ON 07 OCT 2003
 L32 8 SEA FILE=HCAPLUS ABB=ON PLU=ON (PLASMODIUM OR FALCIPARUM) AND ADENOSINETRIPHOSPHATASE
 L33 0 SEA FILE=HCAPLUS ABB=ON PLU=ON L32 AND (MICROTUB? OR MICRO TUBUL? OR ((PROTEIN OR PEPTIDE OR POLYPROTEIN OR POLYPEPTIDE) AND STIMUL?))
 (FILE 'HCAPLUS, MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH, JICST-EPLUS, JAPIO' ENTERED AT 14:47:45 ON 07 OCT 2003)
 L34 1 S L33
 L35 0 S L34 NOT L15

FILE 'HOME' ENTERED AT 14:50:06 ON 07 OCT 2003

THIS PAGE BLANK (USPTO)

10/006780

(FILE 'HCAPLUS, MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH, JICST-EPLUS, JAPIO' ENTERED AT 14:40:14 ON 07 OCT 2003)

L17 131 S "SAKOWICZ R"?/AU
L18 561 S "BERAUD C"?/AU
L19 10837 S "GUO J"?/AU
L20 4357 S "FREEDMAN R"?/AU
L21 3 S L17 AND L18 AND L19 AND L20
L22 42 S L17 AND (L18 OR L19 OR L20)
L23 108 S L18 AND (L19 OR L20)
L24 6 S L19 AND L20
L25 15730 S L17 OR L18 OR L19 OR L20
L26 14 S (PLASMODIUM OR FALCIPARUM) AND (L22 OR L23 OR L25)
L27 18 S L21 OR L24 OR L26
L28 10 DUP REM L27 (8 DUPLICATES REMOVED)

Author(s)

L28 ANSWER 1 OF 10 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 1

ACCESSION NUMBER: 2003:454460 HCAPLUS

DOCUMENT NUMBER: 139:31820

TITLE: **Plasmodium falciparum**
kinesin motor protein KinI-1 with
microtubule-stimulated ATPase activity and uses
for diagnosis and treatment of malaria

INVENTOR(S): **Sakowicz, Roman; Beraud,**
Christophe; Guo, Jun;
Freedman, Richard

PATENT ASSIGNEE(S): Cytokinetics, Inc., USA

SOURCE: PCT Int. Appl., 68 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 2

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2003048320	A2	20030612	WO 2002-US38360	20021127
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG			
US 2003104496	A1	20030605	US 2001-6780	20011130
PRIORITY APPLN. INFO.:			US 2001-6780	A2 20011130
			US 2002-86935	A2 20020228

AB. The present invention relates to protein and cDNA sequences of a new kinesin motor protein, **P. falciparum** KinI-1 (PfKinI-1), and the use of these compns. for the diagnosis, treatment, or prevention of malaria. The invention provides isolated nucleic acid and amino acid sequences of kinesin superfamily motor protein KinI-1, which has microtubule stimulated ATPase activity and/or depolymerizes microtubules. The invention further relates to antibodies to PfKinI-1, methods of screening for PfKinI-1 modulators

THIS PAGE BLANK (USPTO)

using biol. active PfKinI-1, and kits for screening for PfKinI-1 modulators.

L28 ANSWER 2 OF 10 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2003:435206 HCAPLUS

DOCUMENT NUMBER: 139:909

TITLE: Sequences of **Plasmodium falciparum** kinesin KinI-1 and use for treating malaria

INVENTOR(S): Sakowicz, Roman; Beraud, Christophe; Guo, Jun; Freedman, Richard

PATENT ASSIGNEE(S): Cytokinetics, Inc., USA

SOURCE: U.S. Pat. Appl. Publ., 32 pp.

CODEN: USXXCO

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 2

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2003104496	A1	20030605	US 2001-6780	20011130
WO 2003048320	A2	20030612	WO 2002-US38360	20021127
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG				

PRIORITY APPLN. INFO.: US 2001-6780 A2 20011130
US 2002-86935 A2 20020228

AB The invention provides sequences of **Plasmodium falciparum** kinesin KinI-1. The invention also relates to the use of KinI-1 for the diagnosis, treatment, or prevention of malaria.

L28 ANSWER 3 OF 10 WPIDS COPYRIGHT 2003 THOMSON DERWENT on STN

ACCESSION NUMBER: 2003-492112 [46] WPIDS

CROSS REFERENCE: 2002-352124 [38]; 2003-491700 [46]

DOC. NO. CPI: C2003-131606

TITLE: New human microtubule motor protein (kinesin motor protein KinI-3) and nucleic acid, useful for diagnosing, preventing or treating e.g. cancer, restenosis, inflammation, neurological disorders or disorders of vesicular transport.

DERWENT CLASS: B04 D16

INVENTOR(S): BERAUD, C; DAVIES, K A; FREEDMAN, R; GUO, J; PATEL, U A

PATENT ASSIGNEE(S): (CYTO-N) CYTOKINETICS

COUNTRY COUNT: 1

PATENT INFORMATION:

THIS PAGE BLANK (USPTO)

10/006780

PATENT NO	KIND	DATE	WEEK	LA	PG
US 2003036075	A1	20030220	(200346)*		31

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
US 2003036075	A1 CIP of	US 2000-675227	20000929
	Div ex	US 2001-967908	20010928
		US 2002-159151	20020531

FILING DETAILS:

PATENT NO	KIND	PATENT NO
US 2003036075	A1 CIP of	US 6461855

PRIORITY APPLN. INFO: US 2001-967908 20010928; US 2000-675227
20000929; US 2002-159151 20020531

AN 2003-492112 [46] WPIDS
CR 2002-352124 [38]; 2003-491700 [46]
AB US2003036075 A UPAB: 20030719

NOVELTY - An isolated microtubule motor protein, which has greater than 70 % sequence identity to a 1368 amino acid sequence (P1), given in the specification as measured using a sequence comparison algorithm, is new.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

- (1) an isolated nucleic acid sequence:
 - (a) encoding the microtubule motor protein described above, where the motor the protein's activity includes microtubule depolymerization activity; or
 - (b) comprising a sequence which has greater than 60 % sequence identity with nucleotide comprising 6409 base pairs (N1), given in the specification;
- (2) an expression vector comprising the nucleic acid encoding the microtubule motor protein;
- (3) a host cell transfected with the vector of (2);
- (4) screening for modulators of KinI-3 comprising:
 - (a) providing the biologically active KinI-3;
 - (b) contacting the biologically active KinI-3 with a candidate agent in a test and control concentration; and
 - (c) assaying for the level of KinI-3 activity, where the KinI-3 activity consists of binding activity or ATPase activity, and where a change in activity between the test and control concentration indicates a modulator; and
- (5) a compound that modulates KinI-3, which is identified using the method of (4).

ACTIVITY - Cytostatic; Vasotropic; Immunomodulator; Antiinflammatory; Vulnerary; Antirheumatic; Antiarthritic; Antiarteriosclerotic; Antigout; Antipsoriatic; Antidiabetic; Ophthalmological; Immunosuppressive; Neuroprotective.

No biological data is given.

MECHANISM OF ACTION - ATPase Modulator; Microtubule Motor Protein Modulator.

USE - The KinI-3 protein, nucleic acid, or its modulator, is useful for diagnosing, preventing or treating cellular proliferation

THIS PAGE BLANK (USPTO)

(e.g. cancers (e.g. bronchogenic, carcinoma, Kaposi's sarcoma, lymphoma, leukemia, osteoid osteoma, glioblastoma, etc.), hyperplasia, restenosis, cardiac hypertrophy, immune disorders or inflammation), neurological disorders, or disorders of vesicular transport. These disorders also include atherosclerosis, hemangiomas, acoustic neuromas, vascular malfunctions, abnormal wound healing, rheumatoid arthritis, Bechet's disease, gout, psoriasis, diabetic retinopathy, corneal graft rejection, glaucoma, Osler Webber syndrome, etc.. The protein, nucleic acid or the KinI-3 modulator regulates cell cycle, as well as cellular proliferation. The KinI-3 protein or nucleic acid is also useful for screening therapeutic agents or KinI-3 modulators, which may be used for treating the above-mentioned diseases or disorders.

Dwg.0/9

L28 ANSWER 4 OF 10 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 2
 ACCESSION NUMBER: 2002:256423 HCAPLUS
 DOCUMENT NUMBER: 136:274347
 TITLE: Human kinesin motor protein KinI-3, protein and cDNA sequences and tissue expression
 INVENTOR(S): Beraud, Christophe; **Guo, Jun**; **Freedman, Richard**; Patel, Umesh A.; Davies, Katherine A.
 PATENT ASSIGNEE(S): Cytokinetics, Inc., USA
 SOURCE: PCT Int. Appl., 68 pp.
 CODEN: PIXXD2
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 2
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002026929	A2	20020404	WO 2001-US30750	20010928
WO 2002026929	A3	20020613		
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG			
US 6461855	B1	20021008	US 2000-675227	20000929
US 6432659	B1	20020813	US 2000-724510	20001127
US 6436686	B1	20020820	US 2000-723216	20001127
EP 1330535	A2	20030730	EP 2001-977340	20010928
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL, TR			
PRIORITY APPLN. INFO.:			US 2000-675227 A	20000929
			WO 2001-US30750 W	20010928

AB The invention provides isolated nucleic acid and amino acid sequences for human kinesin KinI-3. The invention also relates to antibodies to KinI-3. The invention also relates to methods of screening for KinI-3 modulators using biol. active KinI-3, and kits

THIS PAGE BLANK (USPTO)

for screening for KinI-3 modulators. The invention shows that kinesin KinI-3 is unregulated in lung, colon and breast cancers.

L28 ANSWER 5 OF 10 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 3
 ACCESSION NUMBER: 2002:351338 HCAPLUS
 DOCUMENT NUMBER: 137:74991
 TITLE: A mechanism for microtubule depolymerization by KinI kinesins
 AUTHOR(S): Moores, Carolyn A.; Yu, Ming; Guo, Jun; Beraud, Christophe; Sakowicz, Roman; Milligan, Ronald A.
 CORPORATE SOURCE: Department of Cell Biology, CB227, The Scripps Research Institute, La Jolla, CA, 92037, USA
 SOURCE: Molecular Cell (2002), 9(4), 903-909
 CODEN: MOCEFL; ISSN: 1097-2765
 PUBLISHER: Cell Press
 DOCUMENT TYPE: Journal
 LANGUAGE: English

AB Whereas most kinesins motor along microtubules, KinI kinesins are microtubule depolymerizing machines. Surprisingly, we found that a KinI fragment consisting of only the motor core is capable of ATP-dependent depolymerization. The motor binds along microtubules in all nucleotide states, but in the presence of AMPPNP, microtubule depolymerization also occurs. Structural characterization of the products of AMPPNP-induced destabilization revealed a snapshot of the disassembly machine in action as it precisely deformed a tubulin dimer. While conventional kinesins use the energy of ATP binding to execute a "power-stroke," KinIs use it to bend the underlying protofilament. Thus, the relatively small class-specific differences within the KinI motor core modulate a fundamentally conserved mode of interaction with microtubules to produce a unique depolymerizing activity.

REFERENCE COUNT: 29 THERE ARE 29 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L28 ANSWER 6 OF 10 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC. on STN
 ACCESSION NUMBER: 2003:156397 BIOSIS
 DOCUMENT NUMBER: PREV200300156397
 TITLE: Interaction of a KinI motor core with microtubules.
 AUTHOR(S): Moores, C. A. (1); Hekmat-Nejad, M.; Yu, M.; Sakowicz, R.; Milligan, R. A. (1)
 CORPORATE SOURCE: (1) Dept. of Cell Biology, Scripps Research Institute, La Jolla, CA, USA USA
 SOURCE: Molecular Biology of the Cell, (Nov. 2002, 2002) Vol. 13, No. Supplement, pp. 280a. print.
 Meeting Info.: 42nd Annual Meeting of the American Society for Cell Biology San Francisco, CA, USA December 14-18, 2002 American Society for Cell Biology
 . ISSN: 1059-1524.
 DOCUMENT TYPE: Conference
 LANGUAGE: English

L28 ANSWER 7 OF 10 SCISEARCH COPYRIGHT 2003 THOMSON ISI on STN
 ACCESSION NUMBER: 2002:563016 SCISEARCH
 THE GENUINE ARTICLE: 570RY

THIS PAGE BLANK (USPTO)

TITLE: Potential long-term toxicity of repeated orally administered doses of artemether in rats.

AUTHOR: Xiao S H; Yang Y Q; You Q Q; Utzinger J; Guo H F; Jiao P Y; Mei J Y; **Guo J**; Bergquist R; Tanner M (Reprint)

CORPORATE SOURCE: Swiss Trop Inst, POB, CH-4002 Basel, Switzerland (Reprint); Swiss Trop Inst, CH-4002 Basel, Switzerland; Chinese Acad Prevent Med, Inst Parasit Dis, Shanghai 200025, Peoples R China; Princeton Univ, Off Populat Res, Princeton, NJ 08544 USA; WHO, WHO Special Programme Res & Training Trop Dis, World Bank, UNDP, CH-1211 Geneva 27, Switzerland

COUNTRY OF AUTHOR: Switzerland; Peoples R China; USA

SOURCE: AMERICAN JOURNAL OF TROPICAL MEDICINE AND HYGIENE, (JAN 2002) Vol. 66, No. 1, pp. 30-34. Publisher: AMER SOC TROP MED & HYGIENE, 8000 WESTPARK DR, STE 130, MCLEAN, VA 22101 USA. ISSN: 0002-9637.

DOCUMENT TYPE: Article; Journal

LANGUAGE: English

REFERENCE COUNT: 29

ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS

AB Artemether, an efficacious antimalarial drug, effectively prevents patent schistosome infections and morbidity, as established in laboratory models and in clinical trials. In view of concern about the potential long-term toxicity, rats were treated orally with 80 mg/kg artemether once every 2 weeks for 5 months. After the final treatment, routine blood test results were normal except for reversible reductions of reticulocyte counts and reversible increases in hemoglobin levels. Liver and kidney function and histopathological examination showed no differences between treated and untreated rats. Administration of 400 mg/kg artemether resulted in transient focal vesicle degeneration of the liver or slight damage to the proprius layer lamina of intestinal villi. No damage to the central nervous system tissues, including cerebrum, cerebellum, midbrain, thalamus, pons, medulla oblongata, and spinal cord, was seen at either concentration. There were no alterations in electrocardiograms during the 6-month treatment period. We conclude that 80 mg/kg artemether administered once every 2 weeks is safe, and doses of 400 mg/kg do not result in evidence of neurotoxicology.

L28 ANSWER 8 OF 10 SCISEARCH COPYRIGHT 2003 THOMSON ISI on STN

ACCESSION NUMBER: 2000:878709 SCISEARCH

THE GENUINE ARTICLE: 374GN

TITLE: A statistical-based approach to assessing the fidelity of combinatorial libraries encoded with electrophoretic molecular tags. Development and application of tag decode-assisted single bead LC/MS analysis

AUTHOR: Dolle R E (Reprint); **Guo J**; OBrien L; Jin Y; Piznik M; Bowman K J; Li W N; Egan W J; Cavallaro C L; Roughton A L; Zhao Q; Reader J C; Orlowski M; JacobSamuel B; Carroll C D

CORPORATE SOURCE: PHARMACOEPIA INC, DEPT CHEM, POB 5350, PRINCETON, NJ 08543 (Reprint); PHARMACOEPIA INC, DEPT ANALYT CHEM, PRINCETON, NJ 08543; PHARMACOEPIA INC, DEPT BIOL, PRINCETON, NJ 08543; PHARMACOEPIA INC, CTR INFORMAT & DRUG DISCOVERY, PRINCETON, NJ 08543

THIS PAGE BLANK (USPTO)

10/006780

COUNTRY OF AUTHOR: USA
SOURCE: JOURNAL OF COMBINATORIAL CHEMISTRY, (NOV-DEC 2000)
Vol. 2, No. 6, pp. 716-731.
Publisher: AMER CHEMICAL SOC, 1155 16TH ST, NW,
WASHINGTON, DC 20036.
ISSN: 1520-4766.
DOCUMENT TYPE: Article; Journal
FILE SEGMENT: PHYS; LIFE
LANGUAGE: English
REFERENCE COUNT: 78

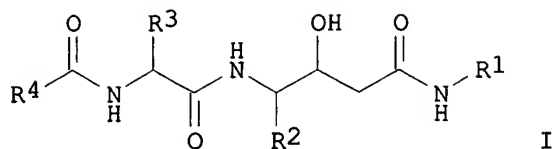
ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS

AB A statistical sampling protocol is described to assess the fidelity of libraries encoded with molecular tags. The methodology, termed library QA, is based on the combined application of tag decode analysis and single bead LC/MS. The physical existence of library compounds eluted from beads is established by comparing the molecular weight predicted by tag decode with empirical measurement. The goal of sampling is to provide information on overall library fidelity and an indication of the performance of individual library synthons. The minimal sampling size n for library QA is $10 \times$ the largest synthon set. Data are reported as proportion (p) \pm lower and upper boundary (lb-ub) computed at the 95% confidence level ($\alpha = 0.05$). As a practical demonstration, library QA was performed on a 25 200-member library of statine amides (size = $40 \times 63 \times 10$). Sampling was conducted three times at n similar to 630 beads per run for a total of 1902 beads. The overall proportions found for the three runs were consistent with one another: $p = 84.4\%$, lb-ub = 81.5-87.2%; $p = 83.1\%$, lb-ub = 80.2-85.95; and $p = 84.5\%$, lb-ub = 81.8-87.3%, suggesting the true value of p is close to 84% compound confirmation. The performance p_i of individual synthons was also computed. Corroboration of QA data with biological screening results obtained from assaying the library against cathepsin D and plasmepsin II is discussed.

L28 ANSWER 9 OF 10 HCAPLUS COPYRIGHT 2003 ACS on STN DUPLICATE 4
ACCESSION NUMBER: 1998:630381 HCAPLUS
DOCUMENT NUMBER: 129:331033
TITLE: Identification of potent inhibitors of
Plasmodium falciparum
plasmepsin II from an encoded statine
combinatorial library
AUTHOR(S): Carroll, Carolyn DiIanni; Patel, Hitesh;
Johnson, Theodore O.; Guo, Tao; Orlowski, Marc;
He, Zhen-Min; Cavallaro, Cullen L.; **Guo,**
Joan; Oksman, Anna; Gluzman, Ilya Y.;
Connelly, James; Chelsky, Daniel; Goldberg,
Daniel E.; Dolle, Roland E.
CORPORATE SOURCE: Department Biology, Pharmacopeia, Inc.,
Princeton, NJ, 08540, USA
SOURCE: Bioorganic & Medicinal Chemistry Letters (1998),
8(17), 2315-2320
CODEN: BMCLE8; ISSN: 0960-894X
PUBLISHER: Elsevier Science Ltd.
DOCUMENT TYPE: Journal
LANGUAGE: English
GI

THIS PAGE BLANK (USPTO)

10/006780



AB An encoded 13,020-member combinatorial library I [R1 = Me, Bu, MeO(CH2)3, Ph(CH2)4, PhCH2, 2-pyridylmethyl, 3-pyridylmethyl; R2 = CH2Ph, Me, CH2CHMe2; R3 = side chain from 31 natural and unnatural L- and D-amino acids; R4 = Ph, 2-naphthylmethyl, 2,4-(MeO)2C6H3, cyclopropyl, 2-carboxy-4-methylcyclohexyl, 4-(4-HOC6H4)C6H4, 5-phenyl-2-pyrrolylethyl, Me3C, 4-methylcyclohexyl, 2-naphthyloxymethyl, Et2NCH2CH2, 4-(Me2NCH2)C6H4, n-heptyl, 3-tetrahydrofuryl, 4-[(HO2C)2CHO]C6H4, HOCHMeCH2, 2,4-Cl2C6H4O(CH2)3, PhCH2] was synthesized containing a statine core. Evaluation of this library with plasmepsin II, an aspartyl protease required for Hb metabolism in the malaria parasite, led to the identification of potent and selective inhibitors as well as novel structure-activity relationships.

REFERENCE COUNT: 18 THERE ARE 18 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L28 ANSWER 10 OF 10 MEDLINE on STN
ACCESSION NUMBER: 93338845 MEDLINE
DOCUMENT NUMBER: 93338845 PubMed ID: 1307277
TITLE: Effect of sodium artesunate on **Plasmodium** yoelii analysed by flow cytometry.
AUTHOR: Han W S; Zhou W Q; Wang B X; Li Z M; Zuo L F; Guo J W
CORPORATE SOURCE: Department of Parasitology, Hebei Medical College, Shijiazhuang.
SOURCE: CHUNG-KUO CHI SHENG CHUNG HSUEH YU CHI SHENG CHUNG PING TSA CHIH CHINESE JOURNAL OF PARASITOLOGY AND PARASITIC DISEASES, (1992) 10 (3) 204-7.
Journal code: 8709992. ISSN: 1000-7423.
PUB. COUNTRY: China
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: Chinese
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199309
ENTRY DATE: Entered STN: 19930917
Last Updated on STN: 19930917
Entered Medline: 19930902

AB The effect of sodium artesunate on **Plasmodium** yoelii-infected mouse erythrocytes was analysed by flow cytometry. The results showed that malarial DNA content in experimental group was obviously decreased 2-5 hours after the drug was administered, fluorescence distribution of malarial DNA almost disappeared within 24 hours after the administration. Thus we deem that sodium artesunate can inhibit the DNA synthesis in P. yoelii.

FILE 'HOME' ENTERED AT 14:43:32 ON 07 OCT 2003

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 2, 2003, 16:13:56 ; Search time 130 Seconds
(without alignments)
2556.706 Million cell updates/sec

Title: US-10-006-780-2

Perfect score: 6842
Sequence: 1 MNSKIVVVRKRLPLELKKK.....KLVQDNKSNMDNNHKK 1288

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virini:*
16: sp_bacteriophage:*
17: sp_archaeoprotein:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6842	100.0	1351	0814Y0	0814Y0 plasmodium
2	903.5	13.2	2577	0814M1	0814M1 plasmodium
3	876.5	12.8	3207	081H00	081H00 plasmodium
4	857	12.5	3248	081C09	081C09 plasmodium
5	854.5	12.5	2235	081B09	081B09 plasmodium
6	851	12.4	5561	081IC9	081IC9 plasmodium
7	850.5	12.4	1422	08MMY3	08MMY3 dictyostel
8	847.5	12.4	1681	081548	081548 plasmodium
9	839	12.3	3178	0812P8	0812P8 plasmodium
10	838	12.2	3256	08MMY5	08MMY5 dictyostel
11	830.5	12.1	3069	081DV7	081DV7 plasmodium
12	830.5	12.1	2874	081C30	081C30 plasmodium
13	830	12.1	2432	081C60	081C60 plasmodium
14	827.5	12.1	2432	081519	081519 plasmodium
15	826.5	12.1	1824	0815A3	0815A3 plasmodium
16	826.5	12.1	2672	081EB4	081EB4 plasmodium

17	824	12.0	901	5	081B06	081B06 dictyostel
18	819.5	12.0	1483	5	081I07	081I07 plasmodium
19	819.5	12.0	4466	5	081LS4	081LS4 plasmodium
20	817.5	11.9	2275	5	081H08	081H08 plasmodium
21	816	11.9	2041	5	081B06	081B06 plasmodium
22	815.5	11.9	1774	5	081A08	081A08 plasmodium
23	813	11.9	3848	5	081D00	081D00 plasmodium
24	810.5	11.8	2472	5	081D05	081D05 dictyostel
25	810	11.8	1200	5	097277	097277 plasmodium
26	808	11.8	2612	5	0815X5	0815X5 plasmodium
27	807	11.8	1918	5	08SSW3	08SSW3 dictyostel
28	804	11.8	2206	5	081B08	081B08 plasmodium
29	804	11.8	8591	5	081B94	081B94 plasmodium
30	803	11.7	1785	5	081520	081520 plasmodium
31	802	11.7	2763	5	081I48	081I48 plasmodium
32	801	11.7	1770	5	081I33	081I33 plasmodium
33	800	11.7	2940	5	081HP9	081HP9 plasmodium
34	799.5	11.7	3569	5	081B08	081B08 plasmodium
35	798	11.7	1605	5	081A05	081A05 plasmodium
36	797.5	11.7	1922	5	081B84	081B84 plasmodium
37	796	11.6	800	10	094G01	094G01 oryza sativ
38	795.5	11.6	706	10	091W81	091W81 arabidopsis
39	795.5	11.6	1182	5	081E15	081E15 plasmodium
40	795	11.6	1437	5	081C24	081C24 plasmodium
41	794	11.6	684	10	0940Y8	0940Y8 arabidopsis
42	793.5	11.6	2849	5	081HY4	081HY4 plasmodium
43	791	11.6	3334	5	081IS9	081IS9 plasmodium
44	790.5	11.6	703	10	09AVP4	09AVP4 nicotiana t
45	789.5	11.5	799	10	091J01	091J01 arabidopsis

ALIGNMENTS

RESULT 1

0814Y0 PRELIMINARY; PRT; 1351 AA.

AC 0814Y0;
DT 01-MAR-2003 (T-REMBLrel. 23, Created)
DT 01-MAR-2003 (T-REMBLrel. 23, Last sequence update)
DE Kinesin-like protein, putative.
GN PFL2165W.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
CX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung B., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pereira M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrett B.;
RA "Genome sequence of the human malaria parasite Plasmodium
falciparum".
RL Nature 419:498-511(2002).
RL EMBL: A014851; AAN36517.1;
SQ SEQUENCE 1351 AA; 156007 MW; 8869594BFD73C788 CRC64;

Query Match 100.0%; Score 6842; DB:5; Length 1351;
Best Local Similarity 100.0%; Pred. No. 7, 1e-268;
Matches 1288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNSKIVVVRKRLPLELKKKDSIIITVANNCTLYIDPRYVDMTKYIERHEFTVDKY 60
DB 64 MNSKIVVVRKRLPLELKKKDSIIITVANNCTLYIDPRYVDMTKYIERHEFTVDKY 123

61 FDDTVADNFTYENTIKPLIIDLYENGCVSCFAYGOTSGKTYTMLGSOPTGSPGIF 120
124 FDDTVADNFTYENTIKPLIIDLYENGCVSCFAYGOTSGKTYTMLGSOPTGSPGIF 183
121 QYAAGDIFFTPIINYDKONTKGIIFISFYEIYCGKLYDLQKRMVAALENGKEVVKOLK 180
184 QYAAGDIFFTPIINYDKONTKGIIFISFYEIYCGKLYDLQKRMVAALENGKEVVKOLK 243
181 ILRVITKEELIKMIDGVLARKIGVSONDESSRSHALITDLKQINNTSLGKIAFDL 240
244 ILRVITKEELIKMIDGVLARKIGVSONDESSRSHALITDLKQINNTSLGKIAFDL 303
241 AGSERGADTVSONKOTOTDGANINRSLALKECIRAMPDSKNHIFPRDSELTGVARDIFV 300
304 AGSERGADTVSONKOTOTDGANINRSLALKECIRAMPDSKNHIFPRDSELTGVARDIFV 363
301 GKSXKIMIANISPTISCEQTLNLTLYSSRVKFNKSTCINEEDTTERISILDSKGS 360
364 GKSXKIMIANISPTISCEQTLNLTLYSSRVKFNKSTCINEEDTTERISILDSKGS 423
361 EMNASSIENVVYIKSNHLLSNNNKXINGKINDKIERNNILKNKSPDKPRGFTSTFGKY 420
424 EMNASSIENVVYIKSNHLLSNNNKXINGKINDKIERNNILKNKSPDKPRGFTSTFGKY 483
421 SSLNDIDKIKKXKGGKJLYKSTLYNDNTINKGNNNNNNNNNDNNNNNNNNNNSS 480
484 SSLNDIDKIKKXKGGKJLYKSTLYNDNTINKGNNNNNNNNNDNNNNNNNNNNSS 543
481 SMVNNMMINMMINNN 540
544 SMVNNMMINMMINNN 603
541 HLNNNDKSIPLHKKKJLRLNKLKXSSCDNTMKNKKNNLHLARHVSGLTMFSDVDPKN 600
604 HLNNNDKSIPLHKKKJLRLNKLKXSSCDNTMKNKKNNLHLARHVSGLTMFSDVDPKN 663
601 KONTEPKSNINMEEDTPKDIYESRNVSNMNGVLLGLKNTKTHDIISTKDENNHNDKIN 660
664 KONTEPKSNINMEEDTPKDIYESRNVSNMNGVLLGLKNTKTHDIISTKDENNHNDKIN 723
661 NGVINIINNSVNSINNSNNSINNSNNSINNSNNSINNSNNSINNSNNSINNSNNS 720
724 NGVINIINNSVNSINNSNNSINNSNNSINNSNNSINNSNNSINNSNNSINNSNNS 783
721 NNNDIIFPAISCDNNMYPITNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 780
784 NNNDIIFPAISCDNNMYPITNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 843
781 HNLFPDPNNKNTSNIONINTKQNDGVNVYSMNFCNYLNDKNTLIDLANKKQKQKIH 840
844 HNLFPDPNNKNTSNIONINTKQNDGVNVYSMNFCNYLNDKNTLIDLANKKQKQKIH 903
841 GCDNNIIONRNDPEKKKKTNFYNNNNIYIVNNNMGNNSPRMKYGLCSGHTSIDMKKNE 900
904 GCDNNIIONRNDPEKKKKTNFYNNNNIYIVNNNMGNNSPRMKYGLCSGHTSIDMKKNE 963
901 MKNNMKNMKNMKNHIIKSNNNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 960
964 MKNNMKNMKNMKNHIIKSNNNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 1023
961 NNINISNTYQNDIITYTINSLDYMSNTLHFKKCYTPTLSTNEDIYKMEGKHRLD 1020
1024 NTNINSNTYQNDIITYTINSLDYMSNTLHFKKCYTPTLSTNEDIYKMEGKHRLD 1083
1021 DDDKYDDNDNNVNDNNNNKNNVNDNNVNDNNVNDNNVNDNNVNDNNVNDNNVNDNN 1080
1084 DDDKYDDNDNNVNDNNNNKNNVNDNNVNDNNVNDNNVNDNNVNDNNVNDNNVNDNN 1143
1081 NFNNNEVLYSFOKAVDTIINNCLNSLSDISMYDDTKELIANLILSKYAEKKNVILKKTIN 1140
1144 NFNNNEVLYSFOKAVDTIINNCLNSLSDISMYDDTKELIANLILSKYAEKKNVILKKTIN 1203
1141 EDIKNMSLEIDKTAQSIYERKRVLLTKLLLFKKNVDTQINNETSDLRKDLVWCHI 1200

1204 EDIKNMSLEIDKTAQSIYERKRVLLTKLLLFKKNVDTQINNETSDLRKDLVWCHI 1263
1201 NPDDQHFHYASRLKIKITINLMIRQWCSSENRLIYOLVYEVOKNSANSVILNVSNN 1260
1264 NPDDQHFHYASRLKIKITINLMIRQWCSSENRLIYOLVYEVOKNSANSVILNVSNN 1323
1261 NGDIITLANKKLVODNIRKSMDDHNNHKK 1288
1324 NGDIITLANKKLVODNIRKSMDDHNNHKK 1351
RESULT 2
081AM1 PRELIMINARY; PRT; 2577 AA.
ID 081AM1
AC 081AM1;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein.
GN M18P1.153
OS Plasmidium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmidium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Seeger K., Murphy L., Harris D., Berriman M., Pain A., Hall N.,
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844507; CAD51342.1;
KW Hypothetical protein.
SQ
SEQUENCE 2577 AA; 299292 MW; 025BBC240587F4D7 CRC64;
Query Match 13.2%; Score 903.5; DB 5; Length 2577;
Best Local Similarity 23.1%; Pred. No. 1.1e-26;
Matches 430; Conservative 228; Mismatches 482; Indels 721; Gaps 89;
19 KKKKOSDIYKNNCTLYIDEPRYKVMTKYIER---HEFTYDKVFDTV--DNFTYV- 71
111 KKKKSKRYCOLND-----KLNFAVFNKGRKCTNSYNNRYNDYKQDNFSGYL 159
72 --ENTIKPLIIDLYENGCVSCFAYGOTSGKTYTMLGSOPTGPIFYAAGDIFT 129
160 SSKNKK-----NSNLKSSSNLFISEPOGQGEPLFSEODIIX 200
130 PLNTY-----DKONTKGIIFISFYEIYCGKLY-DLQKRMVAAL 167
201 LIYFMENFIKFKKENRQNGKQDGRTN-----EYVDMYTKKSYSDKNVKK- 249
168 ENGKEVVKOLKILRVLT-----EELIKMID--GVLLR 201
250 KEGKINIKYKNIQGMVYINKVELAILGRDLSLSLPEVVGENDQYIKEIDEKGRD 309
202 KIG--VNSQDESSRSHALINIDLK-----INKTSLGKIAFDLAGSERGADTVSONKQ 255
310 KIGTVLADENVKRLMLGLVLCNEDEBEVINKVELLLK-----ENNINDIMKQDV 360
256 TOTGAN-----INRSLALKECIRAMPDSKNHIFR----- 287
361 YKTNMSDNKLMENIENDINLTSICYEKKNSSSPSDVTHINIEYEGKIPNNGDTIDEN 420
288 ---DSELTGVARD---IFVGSKSIIM--IANISPTIS-----CCEQT----- 321
421 IIDEKENEKTIKDEKEYWEIGNNDIINDCKQNNKKEISMLGKMGKGVSESNNTISEAVKE 480
322 -----INTLYSSRVKFNKSTCINEEDTTERISILDSK- 358
481 EEKLIQIDINYLQKMEFFNNADIFYPFTIKNF--NLCKNKKKQKMKRNTITERRY 537
359 -GSEMASSIENVVYIKSNHLLSNNN--NKIRGKINDK-----IERNN-----ILK 402
538 LGNKNNNININYN--NNINMYNNNNISLNLKVOCNMNSNGVYIISEBSNNNCSKEYIK 594

0Y	403	NK-SFDKPRG-----	FTSFGKSSLANDI-----	X	428		
Db	505	NKIYSLKRLMGLMMLLKIEKKSRRGRPGGASKYKKELGEOKKKDDCMLVJDSKR		654			
0Y	429	IKKKKKGLI-----	NYKSTL-YNDNTINKGNNNNNNNNNDNNNDNNNDNNNDNNND		478		
Db	655	VKSRHKOKL.PDNRRNNNNNNNNKONLDIMNNNNNNNNNNNNNNNNNNNNNNNNNN		714			
0Y	479	-----	-----	478			
Db	715	YVVYGATNIGNGKKEFANCL.PYDNPEYHNVINKOKLIRRIILNVVDTSNMYSNS		774			
0Y	479	-----	SSMYNNMT---NHMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	SHNNL	519		
Db	775	HNGDNFNLKNGEBSFSSSTNTNTVTLKSSNNNNNNNNNNNNNNNNNNNNNNNNNN	SYNNMI	834			
0Y	520	P-----	OPNYAFTDTSDESSLDNMC-----	HLNNDSIFLHKKGLRDN1-----	560		
Db	835	TEVENNEBN-----	DIGNF--YNPKSCVPYUOKTNNN	IKGVODNYQKXNYVNSFNNSQVH	888		
0Y	561	-----	KLKRRSS--CNIIMKKKKNNLHLARHVSGLT.MF-----	593			
Db	889	PPYHHEEYKONSSKYVCNTYTNNYMHNNINNVYA-GSNMNA.CLNENSVCNDECKIN	IN	947			
0Y	594	-----	SYDPOKONTEP-----	KSNIKMDNRP-KOILVESRVSMGNCV-----	L	636	
Db	948	NNNHEENKAIIDVILIRMYNKSHTORM--	NVPTKOLKNE--	VYKTHAVTVCVYISKT	10020		
0Y	637	LGLNKNTHTDISTKDEHNNDKINNGVINI	INNSVNSINNS--	NNSINN-SMNSN	691		
Db	1003	ISLN-NIWLDSIPNQMTNGYEMEDMNVAFKSDENYTNMMGYMDKINNMMNLVT			10616		
0Y	692	SIYSNYSNOSISDVQIRY--	VNEWDTSNK-----	NNDNIFDAISCDNNYPINTN	742		
Db	1062	KPTRKNNKSNKAGDKTKSRDGGKNNDSNSPQIFEDNHDEYKD-----	DDN--	NNNN	11144		
0Y	743	NN			802		
Db	1115	NN			11659		
0Y	803	NNODGNV--	YSMN-----	FCHNLND-----	KNYL-----	IDLNKEOKDN--	838
Db	1170	NMQNDYNNYKDNBEGNMNTFYNNNITNTYLDENYLCHEKONNTYLONEKGNNDAF					12229
0Y	839	-----	IHCDDNIIQNRNDEFEKKKTNFYNNNNIYVN-----	NM	874		
Db	1230	LGPEONCIATIGSCNNNNNDNTSNMKKCLIRNMFJGNVNVNYNITDNDKCLYNYKI					12899
0Y	875	GNNNSPRKTYGLCSHITSIDMKV-NEKKNEMKCN--	EMKDNH-----	915			
Db	1290	INNEKR-----	GNL.CYIKVKKMMNLIHNNNIIIGNTFFERFDDNTYHMSYDQLYNP		13434		
0Y	916	-----	IKSNNSSSSSSSNN--	NIVN--	NINDDTPQNDYCHND	952	
Db	1344	LNBEEEKKKKKKKKKKSPITLINNOGINNDGNINNDGNINNDGNINNDGNINNDGNIN					14030
0Y	953	FTTIRKKNNTINSNIYQ-NDDIYITNSLN-----	DYMS-----	986			
Db	1404	G-NINNDNINNGIIIGETDDPYLTHIINNMMHMKQNSQNIYHANKDYDLNENDSN					1462
0Y	987	N-TLLHPREKTYTPLSTNEDIYN-KEMEGHILRDODKTDN--	DNNVYN-----	10355			
Db	1463	MMLDEKIGYQKTKSKNEGIMGKKRKYORKDNKVADSVFTGNSSEVFILP8LY					15222
0Y	1036	-----	NNKNVNDNVNNDNVNNDNVNNDKNVNDNVNNDVDDDDVDVFNH	KAPNNEX	10877		
Db	1523	TKVENEIDVANNNT-NNINNNNT-NNINNNINNNNNNNNN--	INNNNNNNSNT	1574			
0Y	1088	LSYFOKAVDTIINNCLSLDISMYDT-----	KEILNITLLS	1125			
Db	1575	CTYNNVNTYEVYINDHNDKKKKKKSDKRDNRSSKKKINIEQKLENNNKESPDNNEMN					1634
0Y	1126	KYKAEKDNVYKYYN-----	EDIKNSGLEIBIDKTAQSIYEKR	1163			

[illegible]

Db	218	FGSTSLYNNNNNNNNNNNNNNDKYLNTYYASPRNOQIYLN-FOQINNHHNNNI-----	270
Oy	242	GSEGGADTVSONKQOTQDGANINRSLSLALKECIRAMDSDKXHIFFRDSSELYKLRJDFVG	301
Db	271	-----NSYSNNKCFNTYMLMKNN-----LEETBRYPPNDYRNI-INNNML-----	309
Oy	302	KSKSIMIANISPTISCCQIANTLTRYSRVKNFKONKSTCINEEDYTERISLIDSKGSE	361
Db	310	NSNNYHMMNIYPELN-----NTYNNMMNTYHLNNVHASTSLÖNNYVNNPNFDH-----	356
Oy	362	MNASIEIENVVJKSNHLSNNNNKKINRGKIDKERNNI-----	400
Db	357	-SMKTIKQETNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNIYSPELDTHLRYKONINEYDI	415
Oy	401	-----LNKRSF-----DKPREGFTSF-----	417
Db	416	ELKICTTLANNKYICRYGIGHSGMTLYNHRICICLGAVKEDIDYVLAHFVLNNGNITYGS	475
Oy	418	GKYSLSLNDIDKIKKKKK-----KGLINYKSTLYNDNTINKKHNNNNNNNDNDNDNDNDN	472
Db	476	GKKCSVSIIRKQIOTDRIHITFKHIIKTPLYLXSKEDKONKNNNNNNNNNNNNNNNNNSN	535
Oy	473	NNNNNDSSSMYNNML-NHMINNNIINNINNNNNNNNNNNNNNNNNNNNNNNNNHNLPOPNVAFDT	530
Db	536	NNNNNSNSNSNNNTNSNSNSNSNSNNNTNSNSNSNNNNNNNNNNNNNNNNNNNNNN- -----TTNNNS	588
Oy	531	DFSSLDNNCHLNNNDK-----SIFLHKKLRDNI-----	560
Db	569	SSSNNNSNNNNYHNNYKAEKELANSSSLSEHSITIMANDNINNINNHHNINNTIDLANMNV	648
Oy	561	---KLKONSSCDNINKKKNMLHLARHSVSGKLT-----MFSYDPOK	599
Db	649	NQSNMKEKNNNIIDYNNNNNNNDNYSNHNLNANCINKLYTNNTYFTEDSQKRAPLOTYNTSK	708
Oy	600	NKD-----NTPFKS-----NINMEMENTKOLLYESRANVSNNNGVVLGLNK---NTHHDI	647
Db	709	MTNNFLANNFTSSYSPFINNMDSN-----IYHTTNNPNKOLNNININISINKNNNI	763
Oy	648	---STQDENNDK-----INGVINIINNSVANSIN-----NSNM	681
Db	764	FNMMHNLHLDHASTIQNNLKYGHNNVNTNIIYNPIMANNINNOQINNLSPNNKNEDN	823
Oy	682	SINNSNM--SNSIY-----KSNYSNQSIS--DVOIRYVNE	714
Db	824	EINHDSDDDSNSHITLKSQKQKQFALNPKYQHÖNNHINNINIONNLEQIKEKD	883
Oy	715	MDTSN--KN-----NDNIFPALSICD-----NNMY	737
Db	884	QONNHKEIKONELNDTISIEDTNDNSYKITYSSDISQNTLNSFOHNKEISVPMY	943
Oy	738	PVITNNNNNNNNNNNNNNNNNNNDIVE--NYNNRQGTNSMKLYAYSNHLFOPDNKATSN	795
Db	944	NNIILDNNNNIIDDNNNNNNNYFCICGINTKE-----YKNYINTYVPPNAN-HIY	994
Oy	796	QNIINTKNQDQNVVYSMNFCHYNLNDKUYLI-----DLNKEOK-----	835
Db	995	NNNNISYNSASVNNYVVTYVNNFHSYHNHYLIHNNFHPYNIYDNIQNTQOKKLYINON	1054
Oy	836	DKXIHGCDNNI-----IONRDPFKKKKTINYYNN--NIYVUNNMG--NNNSPRM	882
Db	1055	DERONNSFMHINTDPHKVYVSNNF-LPIITFYNNNLHNNLILTESNNLNRKNEENDITPS	1113
Oy	883	-----KYGCG-----SHTSIDMKONEMKNNEM-----KONEMKDNIKSNNNN- -----922	922
Db	1114	YSOIHNHQICXKVEEITYNSI-NÖNTNPNNNVMMALNTSNPIPLDNTYNSKKIITYK	1172
Oy	923	-----SSSSSSNNNI-YNNIND-DTFON--DYCHANDFTTIRKONTYINSNI-----YON	971
Db	1173	HIINHINQKNNVVEYENLNSODNTÖNKETFCNODLI-----NSSNINNISISYFON	1266
Oy	972	DDIIYF-----INSLNDYMSUTLHFKKITYPPLSTYMEDLYNKEMEG--KHIRLDDQ	1023

```

Db 1227 MNDEFYKKSQMÖNHNDIYKINT--SENVGSPHTNKTSTSYNHKKGEQHEQTEÖNQ 1283
Oy 1024 KYDNDNNNVADNNKKNVNDNVNDN--NVNDNVNDNDKKNVNDNNVNDNDDDVDYFNHK 1080
Db 1284 NNEQNSBÖNEIQEIEQNEIQVAKÖNVAÖNVAÖNEQVEQVÖNEQVAKÖNEQVÖKQAEQ 1343
Oy 1081 NFNNNEX---LSYFOQNVDTIINNCLNSIDISMYDQTEIINNTILSLKYKAEKDNVIK 1136
Db 1344 NSNNESIKNTVVEFFKKNQÖINTS--NNV-ISQÖQHDNTNIIINININ---IKENINR 1396
Oy 1137 KYINEDIKNMSLEEDIKTAOSIYEKKRVLLTYKLLLPFKRVDPQINNE---TSDLRDQV 1193
Db 1397 HKINE-----PÖWESKNKIDIEKKNCLTYK---TDKONDENDENDENTYKANDIV 1445
Oy 1194 MCHICNNNPDDQFHYAVYSRLKEDIIINLIMROIWCESENILRLYOFLVVEY----- 1245
Db 1446 ICNNHNNSHYÖKNXY---NMNESIMENNI---IITEGENIMNSTE---EYFTEMLIK 1496
Oy 1246 ----QNSANSVLLNVSNNQDITILANKLVONIKRSMÖHNHKK 1287
Db 1497 DSELEKNNSDTKFLIKL--NNEIKKEEK--KDNII-NIFINNIIYE 1536

```

RESULT 4					
ID	Q8ICP9	PRELIMINARY;	PRT;	3248 AA.	
AC	Q8ICP9				
DT	01-MAR-2003	(TREMBLrel. 23, Created)			
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)			
DE	Hypothetical protein.				
GN	MAL6P1.307.				
OS	Plasmodium falciparum (isolate 3D7).				
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.				
NCBI_TaxId=36329;					
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Cherevach I., Davis P., Goodhead I., Stevens K., Mungall K.,				
RA	Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,				
RA	Armond D., Sanders M., Hayes R., Hall S., Quail M., Barrett B.,				
RL	Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.				
DR	EMBL; AL684505; CAD50391.1; -				
KW	Hypothetical protein.				
SQ	SEQUENCE 3248 AA; 374892 MW; F72C129A97E38466 CRC64;				
Query Match		12.5%; Score 857; DB 5; Length 3248;			
Best Local Similarity		23.0%; Pred. No. 1e-26;			
Matches 347;	Conservative 220; Mismatches 413; Indels 530; Gaps 62;				
QY	202 KIGVNSQNDES-----SRSHAILN-----IDLKDINKNTSLGKIAFTDLAGSERG	246			
DB	371 KVTIGSYNNEDNKFFNMGMHVIKNEMDGANDANKIKENEDVN-----IPICGSSEKS	420			
QY	247 ADYTSQNKQTOTDGANINRSLALKECIRANDSDSKNHIFPFDSELTKVLRIPIFGSKSI	306			
DB	421 PLNIQNN-----NI-----EINH-----WRDNFL--SNDI	443			
QY	307 MIANISPTISCCEOTLTLTRYSRVKFNPKNS-----TCINEEDDTINTERISILD	356			
DB	444 M-----NTTIPSPPKKIKIEKRMMKIMHPNVYIKMDERNMD--MDGMS	487			
QY	357 SKGSEMNASSTENV-----VIKSNHL	378			
DB	488 ANNIDPMNGSVNNIDPMNGCATNVINDMGSVNNIDINMGCGNNIDINMGCGNNIDMDHM	547			
QY	379 SNNNN--NKN-----RGKINDKIERNNITLKNSFPKPR---EGTSTPGAKS---	421			
DB	548 SNNNDLIMNNKYNIHMNIINNKI PNKFDAENENIRIGKTDHPXYLNGKDYATRSCTNENNI	607			
QY	422 --SLNDIDKIKKKN---KKGLINYKSTLYNDNTINKGHANNNNNNNNNDNDNNNNN	475			
DB	608 EYNNEENDDKGNKKIIRKNNIMENTNNNNNNNDNDSSNNNNNSNNNNNSNNNNIN	667			

[illegible][illegible]

[illegible][illegible]

Dp		1284	NNNNSSGSNCVANTFLIRNNLMK-----NPNQNRNVANNNSINIH7SSSNINNMM-----	1333
Oy		889	VWYSNPFCHYANLDKGYLIDLANKGOK-K-NHGCDDNI-----IQNRN	851
Dd		1335	-NYNHILTHQKJNSNNNNNNNNNNMNGCLYNHQMSNKVHGTPVPIMVHWGGTISININ	1393
Oy		852	DPEKKKTNTFYVNNNNIVIVYNNMG-----NNNSPRMKYGLCSGHTSIDMHKNEMKNEMK	907
Dd		1334	HMMNINNMNMNMNNINNMSSSNEMGSPINHPIPHKGCGTHMSNOITRRAMNMGLSNPINV	1455
Oy		908	DNEMK-----DNHIKSNNNSSSSSSSN--NNIYNNINDDTEQ-----	944
Dd		1454	DNKMNVLRKNSCTVNFPMNPNPMVSMMNNYNNSMNMNNSLSLNNICSGGITIMPENMKGX	1513
Oy		945	-NDYCHDNTFTIRKKNNTNIS-----NIYONDIIYTTINSLDY	984
Dd		1514	IKDTFINNSSF-----RNNNNINNMNMNMNSMNMHMNSMNMHNSMNMHNMNMNMNMNMNM	1569
Oy		985	MSNTL-----LHFKEKT--YP-----TLST	1003
Dd		1570	MNNRSEONIIANNMIKTEMRRNNLMNMKLVRNQHMNMKYVSEKYPMTTKMIQNNNNVT	1629
Oy		1004	NEDI-----YNKMEG-----KHRLDDDKY	1025
Dd		1630	NRDMNIIINNNGINGINNMGFIVNMKANLTINKMKKITENTYTTHLLHNDPLSPMDTY	1688
Oy		1026	D-DNDNNNVNKKNVNDNVNVDNNVDNNDKNVDNNDVNDDDDVDVFHNIKNFNN	1084
Dd		1690	DIGHNMNMI--NNNMNI--NMVMNINIINI--NSINNIINMN-----SIHSV---N	1722
Oy		1085	NEYLSYPOKAVDTIINNCLNSLIDS--MYDDTEILANNILLSKKAEKDVVKYKINED	1142
Dd		1733	TKFMLNKEDDKTLHPNLYLNQMVOSNTYYHOTYEERINTYNTHTSSPTTIMKKKHKN-S	1791
Oy		1143	IKNMSLEID-----KTAOSIYEKKVLTLK-----LLFPKXV--PTQIINE---	1184
Dd		1792	LKNSKEDDLNANNOKRSVKIWMKYKNDLPFKIPLERNYEGFTLGOKNIKDEDTIINDEXTF	1851
Oy		1185	----TSDRKDLVMCHICGNPPDDOFHFVAARSLEKDIINLIMEROI-----WCSE	1232
Dd		1852	MLENNEDKEDDAIYCDLSNIN-----PKTFPOYNDILNKLNLIKLNDKTNFNKNCISO	1905
Oy		1233	MLRLLYQ--ELVVE-----YONKSANSVLLVSSNNG	1262
Dd		1906	----LYBEIFLIEKFLDISHKYHNNETHTFYNNCLKAYIYONKE-NTNLIHIKEENV	1960
Oy		1263	DIILLNKKL-----VQDNIKSMSDHN	1283
Dd		1961	D-----NKNIGHGLDHLTHENNETKEEDH	1985

RESULT 7

OBMMY3	PRELIMINARY;	PRT; 1422 AA.
ID	OBMMY3	
AC	OBMMY3;	
DT	01-OCT-2002 (TREMBREL. 22, Created)	
DT	01-OCT-2002 (TREMBREL. 22, Last sequence update)	
DT	01-MAR-2003 (TREMBREL. 23, Last annotation update)	
DE	CG5792 protein. 6/101.	
OC	Dicystosellium discoideum (Slime mold).	
CC	Eukaryota; Mycetozoa; Dicystosellida; Dicystosellium.	
OX	NCBI_TaxId=44689;	
RN	[1]	
RC	SEQUENCE FROM N.A.	
RP	STRAIN=AX;	
RA	Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,	
RA	Lehmann R., Baumgart C., Paria G., April J.F., Guigo R., Kumpf K.,	
RA	Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.,	
RT	"Sequence and Analysis of Chromosome 2 of Dicystosellium."	
RL	Submitted (May-2002) to the ENBL/Genbank/DBJ databases.	
DR	EMBL, AC116965; AAM3216.1; --	
DR	InterPro, IPR002110; ANK.	
DR	SMART; SMO0248; ANK. 2.	

[illegible]

[illegible]

ID	081548	PRELIMINARY;	PRT; 1681 AA.
AC	081548;		
DT	01-MAR-2003 (TREMBlrel. 23, Created)		
DT	01-MAR-2003 (TREMBlrel. 23, Last sequence update)		
DT	01-MAR-2003 (TREMBlrel. 23, Last annotation update)		
DE	Hypothetical protein.		
GN	PfPR1815C.		
OS	Plasmodium falciparum (isolate 3D7).		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_TaxID=36529;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=3D7;		
RX	MEDLINE=22255705; PubMed=12368864;		
RA	Gardner W.J., Hall N., Fung E., White O., Berriman M., Hyman R.W., James K.,		
RA	Carlton J.M., Patn A., Nelson K.E., Bowman S., Paulsen I.T., Kyes S.,		
RA	Bisanz J.A., Rutherford K., Salberg S.L., Craig A., Iyken S.,		
RA	Chan M.-S., Nene V., Shalloo S.J., Suh B., Peterson J., Angiuoli S.,		
RA	Perera M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,		
RA	Martin D.M.A., Fairlamb A.H., Franholz M.J., Roos D.S., Ralph S.A.,		
RA	McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,		
RA	Venter J.C., Cantucci D.J., Hoffman S.W., Newbold C., Davis R.W.,		
RA	Frazer C.M., Barrett B.;		
RT	"Genome sequence of the human malaria parasite Plasmodium		
RT	falciparum.";		
RL	Nature 419:498-511 (2002).		
DR	EMBL, AB014850, AAN35449.1, -.		
SK	Hypothetical protein.		
SQ	SEQUENCE 1681 AA; 192792 MW; 3985DBE957CBA86 CRC64;		

Query Match	12.4%;	Score 847.5;	DB 5;	Length 1681;
Best Local Similarity	24.8%;	Pred. No. 1.4e-26;		
Matches 329;	Conservative 223;	Mismatches 406;	Indels 369;	Gaps 63

```

QY 170 GKKEVVVVDKILR-----VLTKEELLKIMDG--LRRKIGVNSQNDSSRSLALN 220
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db 3 GDEMKMKQLOLLKEKMMKNAKAÖKSEEVYKSESVSNLNSVNSVNSVAN---VRS 58
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
QY 221 IDLKINKNTSLGKIAFIIDLAGSRGADVNSQNKQQT---DGANINRSILAKECIRAM 277
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db 59 VSKDIFRSVKGSKIFL-----DPSFQKKLMVHPKNGRNI----- 95
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
QY 278 DSDKNHIIPERDESLFVNLADIFWGSKSMTI-----ANISPTISCEQTLNLTLYSRVK 332
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db 96 -----PÖTSKNVS-----FVSKKNSLSLKLTTEELI.FDTPLSSEEVWRRKNEKKKKE 142
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::

```

Qy	333	NEFKRSCINCNEEDJTER:SIIDSQSENNAS----	SIENVI-----	KSHTLLSN	384							
Db	143	LQKNQUTINISE-FINKQSIQNSPKNKSHLEFKI	QYTNLAI	SODQNGTNTNNNN	201							
Qy	385	KINRGKIN-----	DKIERNNILKNKSFDPKREGT-----	STEGKYS	421							
Db	202	NNNNVVNVVAVVDNNTHTNNLQNNSLINGNL	LÖNNNS	SYNVDTI	KVTSILANTIS	FCNF	261					
Qy	422	SLNDIDIKKQKKKGLIN----	YKSTI	YNDTINKGN-----	NNNN--	NNNDNNND	470					
Db	262	TQNISNNEKQCVGMINTPEFPNNSCKDEKIL	TTYS	VNGTVFNINLI	VKDDNDNDND	321						
Qy	471	NNNNNNSSVMVNMIMHINNINNINNINNINN	NNNNNNNNNNNNNNNNNNNNNNNN	SHNL	POPVAF	DTDS	530					
Db	322	EDNNNN-----	NNNDEDN-----	NNNNNNNDEDNNNNNL-----	353							
Qy	531	DFSJLDDMGNLNNNDKSIPLAKKLNLDIKLR	-----	SSCDNIMK-KKNNLHLAR	SHV	586						
Db	354	-----	LN-----	KSTFPND	ISGTAL	KNDYI	FSJDEKYNKDKENNI-----	392				
Qy	587	GSKLTWFSYDPQKKKNTFFKSNINKMEDN----	TPQDIL	YESRNVSMAGVLLGLKNT	643							
Db	393	-----	NESHNIFKNDKNNNGTNI	FTSNFMTSEKIL	ISMKANTII	INNNOQT	439					
Qy	644	HHIDSTD-----	EHNDKINNGVNI	INN-----	SVANSIN	677						
Db	440	QNNMSQSPSTSNNVQVHNHTTSLNGTSAGI	QKVAVDNIC	CPKNGPTRINN	VNNVNN	499						
Qy	678	SNMGSINNSMNSIYKSNV-----	SNQSID	VOIRVNMEDTSN	KNDNI	FF	727					
Db	500	VNNNI	PNNASVNNNGIF	PSGKNL	FCPIARTNNL	FLQKIESNI	INTNKNNNNNNNN	LMFN	559			
Qy	728	DAISCDNNMY--	PVITNN--	NNNNNNNNNNNNNNI	DVENTNN	DGTNSKLYA	YNSHNL	783				
Db	560	KTISDSTNIGKPCITNNI	FNNTLGNH	INKNNNTLE	FNNTINTNN	NNNNI	-----	SNNS	613			
Qy	784	QOPDNKGTSGIQVINTNNKNNQ-----	DGVNYS	SMNCHYNL-----	NDKXYL-ID	LANK	EOJKD	837				
Db	614	FNPLITNNLTLC-----	NEDKVV	SVSNH	ENI	ENI	YOENKSE	ENDENNLN	YSKVENKES	667		
Qy	838	NHCGCDNNIIONR	ADFEKKKKTPEFYNNNNI	VIYNNNGNNNS	PRMKYGL	CGSHTS	IDMK	897				
Db	668	TMTIGDNNNV	FSFNFKANTLT	LTNNNDNNNNNN	DDNNNN--	NY-----	NNN	714				
Qy	898	NNEKKNEMKDNKDHIKSN-----	NSSSSSS	NNNNIYNNI	INDDT	QND	946					
Db	715	NNNNNNNNNDNI	VAGDHLKSNS	ECVOGQNNNEQNN	EMVYKONT	II	INTNNND	NVJKE	774			
Qy	947	YCHADNFTIRKNT-----	NIN-S	NIYQ-----	NDI	IYINS	LNDYMS	TLT	990			
Db	775	II-----	TWESPENNTI	CKGEDI	HIKFOI	NVSNMDE	VAKSGED	QOHYINN-N	NINESKEPK	830		
Qy	991	HEFEKYT-----	XPTIST	NEDIYKMEG	KHILDDOD	KYDNDNN	VNDNN	KNNVNDN	10444			
Db	831	NIKTFTFSYFNFPISITFSB--	NKK-EP	DOQNGN	DDNNDDNN	DDNN	DDNN	DDNN	DDNN	885		
Qy	1045	VDNNNVNDNNVNDKANNVNDNNVDNDD	VDVDFHN-----	IKN-----	FNNE	YLS	FOKX--	1094				
Db	886	ENNEENENN-ENNE	NNKNNENK	NKDYDN	KFNRTGD	IITNDS	LTJN--	LCIFKQNET	942			
Qy	1095	VDTIINCLNCLDLS-----	MYDPT-----	KEILNN	LTLS	1125						
Db	943	VENIPNNNDNNDDSSNNKEEC	LNKKQTL	MYNTIKODG	SGNSVNNNS	QNNNN	LENNLS	1002				
Qy	1126	KYKAEKDNV-----	KKYNEDI	KMSLE	EDI	KAQSI	IYERK	VLLTKL	LLLF	1173		
Db	1003	GNNPNDNII	LASKVEK	PLGSS	EKTKEP	IELMT	PNKIN	ESNS	NFP	FKNNVNTN	IL-----	1055
Qy	1174	KKAVDTQJNNFSDLR-----	KDLV	MCHIC	NNPDDOF	FHFYAV	SRL	EKDII	NLM	ROI	1227	
Db	1060	SKGEBIE-NMENA	ELNFR:PKQGV	NLS	SEYENKE	EPED	KFKI	YITIN-----	1109			
Qy	1228	WCSENLRLYYOFL-----	VVEYON	KANSV	LNV	SSNGDI	IILNK	KVQ	DN	1278		


```

Db 1110 -CINKNFINTYPSNVNINNFEDONSANT--YDALSKTIDMEL--KRMSSDNNNN 1164
QY 1279 SMDNNI 1285
Db 1165 NNNNNNI 1171

RESULT 9
0812P8 PRELIMINARY; FRT: 3178 AA.
ID 0812P8
AC 0812P8;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Protein kinase, putative.
GN PF1280C.
OS Plasmidium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmidium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368687;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Cotton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ommond D., Price C., Quail M.A., Rabinowitch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrett B.G;
RT "Sequence of Plasmidium falciparum chromosomes 1, 3-9 and 13."
RL Nature 419:557-531(2002).
DR EMBL; AL929358; CAD51942.1; -.
KW Kinase.
SQ SEQUENCE 3178 AA; 366959 MW; 82PB35F0914C4DED CRC64;

Query Match 12.3%; Score 839; DB 5; Length 3178;
Best Local Similarity 23.6%; Pred. No. 5.4e-26;
Matches 332; Conservative 210; Mismatches 420; Indels 444; Gaps 55;

QY 196 DGVLLRKIGVNSQND--SSRSHALNIDLDKINKNTSLGKIAFIDLAGEGADTVSQNK 254
Db 207 DFLNNKTI--INSHKEIHTTNDVLNHNIMSSNNISKSNNNYFGNEK---YDKDN 262
QY 255 QTOTDGNIRSLALKECIRPADSDKNIH PFRDSELTQVLRDI FVGKSKSIMANISPT 314
Db 263 YTHHDHFPYSSSYTEGAIKCLKN--NPVILRDS----- 295
QY 315 ISCEQTLNLTFRYSRVKFNKSTCINEEDDNTTERISILDSKSEMN---ASSIENV 371
Db 296 ---CNEINITFEDINSNDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 348
QY 372 IKSN---HLLSNNNNNKIRGKINDIERNNILKNSFD---KPRGFTSTFGKYS-- 421
Db 349 ILEKKLILNLTINNINENNPNNSYDCKGLPESDKDMNGLTTPQEBICVKQDKEHYK 408
QY 422 -SLNDIDK---IKKKKKKGL-----INYSSTLYNDN-----TINKKAN 456
Db 409 DFLNVLNKKEMINFTKDNNSNNINININIDNVNTKSSLFNDKIFMDNHTDNTIMVTNNNNNN 468
QY 457 NNNNNNDNNNDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 488
Db 469 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 528
QY 489 -----HMINNINNNI 499

```

```

Db 529 YKKLTPLKXKNEDEDEDIKLEGDPNDQVLDNVNDDYDDDNQTHRHSKNNNDNNN 588
QY 500 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 556
Db 589 DNNNDNNNNNDNNININDNDNDCGYLYKNSSTLDESLKQEVN--TSIFSTYKAGI 647
QY 557 RDNILKLNSSCDNINAKKKNNLHLARHSVSKLTWFSYDPQKNKNTFFKSKINKEON 616
Db 648 YNNNDVINFPNDNIMEYKN-----ISSLSYKQVCDASTYNNNDIN----- 650
QY 617 TPKDILYESRNVSNMGNVLLGLNKYTHDIDSTFDENNDNKINGVINIINNSNVNSIN 676
Db 691 -----NNTNNTNNTNNTNNTNNTNNTNNTNNTNNTNNTNNTNNTNNTNNTNNT 728
QY 677 NSMNSINNSNNNSNITSYKNSNSOSISDVQRYNEMDTN----- 719
Db 739 NNTYNNIRVSNNNHCDHNSYDTENI IYEDLLYNNDLKXKSDHFIDEGBIREYNYK 798
QY 720 -----KANNDIFPDALSCDNNMY-----PNTNNNNNNNN-----NNNNN 753
Db 799 RLISYOKIQTNESEDIRKNDNYNMY--NLANDMFSSNDENIKTEPNNSISSFYKNTS 856
QY 754 N---NNNIDVENYNNRDTNNSMKLYAYNSHLFQPDNNKNTSNIGINTNKNQDGNVA 810
Db 857 NFLINDIDCDNKKDQPSVEPKRHISYNSYEKEEQNLKHSLSLENNYINVENARINNF 916
QY 811 YSNPCHYLNLDKNYIDL-----NNKEQDKNIGCDNNIIONRDPFKKKK 858
Db 917 FS---CNVYLTDDTKTNNNTTYTNGNNNNNNNNNIDDDKNDGDDKNNNIDLVNRYKRN 973
QY 859 TNYFYNNN---NIVYNNNMGNNSPRMKYGLCSHTSIDNM-----K 897
Db 974 SNSSNNIHYDNKYMPSPYKKNISTINMYEPYSHSGANNVYIDEHAFYGSFLNKEKA 1033
QY 898 NNEKKN-----EKDNEMKDNHISKNNSSSSSSSSNNNY-----N 935
Db 1034 NKKIKHKNNNNNKSYIDHDITVDNFI SDPFNNMKGEDNVYNNVYDEDDVNNNDGN 1093
QY 936 NIN-----DDDFQNDYCH-----NDTFTIRKKNNTINS- 966
Db 1094 NIDGNDIDGNDIDGNDIDGNDIDGNDIDGNDIDGNDIDGNDIDGNDIDGNDIDGNDIDG 1153
QY 967 NIYONDIIYITINSUNDY-----MSNTLHFKKRYTPTSTNEDI---YN 1009
Db 1154 NMSSESVAISYHDLKKNVYKNSINLEQKCKMPSSTHIDICANTYEKINHMDVAVHSN 1213
QY 1010 KMEGGRILD---DQKXDDNDN-----NNVDN-N 1036
Db 1214 NHHKTHTDMKQVFPDQHCYVSGNMEYIKDNNNTVNSINSNMSNMSNLSLNDVNNVNNIN 1273
QY 1037 NKNVNDVNDVNDVNDVNDVNDVNDVNDVNDVNDVNDVNDVNDVNDVNDVNDVNDVNDV 1094
Db 1274 NVNNNM--NVNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1322
QY 1095 VDTIINNCNSLDSMVDYDTEILNINILSKYKAKQNVIKVYINEDIKMSLEEDIKT 1154
Db 1323 ---VHNS-NSMKSGICEYMMNNYMLNIRMS--KQDDNNTKVDYGNHYN----- 1366
QY 1155 AOSIYERKRYLLTKLLLFKKNVDTQINN---ETSDLRKDLVYCHI CNVPPDQFHFYAY 1211
Db 1367 SQIIGMKKQDVLYSDMSNMVEMNNNNINNNYINNNSNNHDDL-----PNNINDQYN- 1417
QY 1212 SRLKCDIT---NLMLKQIWESENLRLLYPLVVEYQNKANSVLANVSSNNGDI--- 1264
Db 1418 ---NDSLKFKKNFTMSNNL--CGSTNMN-----NINILNVNNLKDINTM 1457
QY 1265 ILLANKLVODNITKN---SMD-HNNTH 1286
Db 1458 DLLNSRKATNSMKQVYSTMDILANVH 1483

```

RESULT 10

[illegible]

QY	1272	VDNKNKSM--DHNKI	1285
Db	3077	ISHNKKIMPYINNI	3093
RESULT 12			
ID	081350	PRELIMINARY;	PRT, 3069 AA.
AC	081350		
DT	01-MAR-2003	(TREMBLrel. 23, Created)	
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)	
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)	
DE		Hypothetical protein.	
GN	P10495		
OS		Plasmodium falciparum (isolate 3D7).	
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_Taxid=36329;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22255708; PubMed=12368867;		
RA	Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,		
RA	Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,		
RA	Buckee C.O., Burrows C., Chervach I., Chillingworth C.,		
RA	Chillingworth T., Christodoulou Z., Clark L., Clark R., Corcoran C.,		
RA	Croin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,		
RA	Fellwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,		
RA	Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,		
RA	Humphrey S., Jagels K., James K.D., Johnson D., Kethornou A.,		
RA	Knights A., Kontoforov B., Kyes K., Larke N., Lawson D., Lennard N.,		
RA	Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,		
RA	Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,		
RA	Rajandream M.A., Ruter S., Rutherford K.M., Sanders S., Simmonds M.,		
RA	Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,		
RA	Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,		
RA	Sulston J.E., Craig A., Newbold C., Barrett B.G;		
RT	"Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";		
RL	Nature 419:527-531(2002).		
DR	EMBL; AL929356; CAD51785.1; -		
KW	Hypothetical protein.		
SO	SEQUENCE 3069 AA; 368341 MW; 731D5B1C386643B3 CRC64;		
QY	Query Match	12.1%; Score 830.5; DB 5; Length 3069;	
	Best Local Similarity	21.6%; Pred. No.1.2e-25;	
	Matches 374; Conservative 230; Mismatches 483; Indels 643; Gaps 72		
QY	21 KKSDDITKVNKCTLYIDEPYKVDWTKYERHFIYDKFDFVDYDNFTYENTIKP--77		
Db	614 EKDEKYLNANNNTCSY-EONKKYMD-TNYLYNHE--DK-----NNYSTKMDLNERNK 662		
QY	78 --LIIDYENCVCSCFAYGOTGSGKTYTLMGSPY-QGSDTPGIFQYAADIF--TFL 131		
Db	663 RFSSKNLVNNSHISEC-----PYDEQISKFLFIHMNNNILKDEKTL 704		
QY	132 NIYD-----KDN-----TKGIFISFEYI 149		
Db	705 TIFPDEKSDDDHHHHNNNNNNNNNNINIDNNVLNHNVLNSNNFVSHGAKNGFKNNLNI 764		
QY	150 YCGKLYD--LQKKKVAALENGKKEVVVKDIKILRVLTKEELILMIDGVLTKIGV 205		
Db	765 ECKKINDQSDCLFYDNRCQIMCPYKKDIFNNVDYKTLNLKNDNL-----809		
QY	206 NSQNDSSRSALINIDLKDN-----KNTSLGKIAF-----IDLASRGADIVS 251		
Db	810 NYSSYSNHSIHRKYSDDINSLSIKKISCKISCNNIKGGHMYKNCYNHDMLYNEHEKSM 869		
QY	252 Q-----NKOTQTDGA--NINSLALKECIRAMSDKN-----282		
Db	870 EYERVAKAKNCRRLIQKDNVRSYINIKSIHNEHRIINIYNNNNNCDNNNCDNNRDYHCD 929		
QY	283 --NHPFSDSLTKYLRDIFVFGKSKSMTANISPTISCEQTL-----NLRKSSRYKN 333		
Db	930 TRAHKDSNENK-----YNGNDMI CNDHNLISHTLSKYDNFVLSDDNNNDNNNNSTLKEEN 984		

[illegible]

Db 1962 -----NNNNNNYFFCHANNKTEMEKEITEIQMLSLSKTIRLGNVANDENNTNOPIFNKKFS 2013

Qy 1129 AEK--DNVIAKYIINEDIKNNLSLEEDKTAOSIYEKKRVLLTKLLLFKNVQTOINNET 1185

Db 2014 FEDFOLDNFINSMDLMDKWD--ENIINKMIEL-ERRKYAINCIMIEKLKRD-----QNES 2064

Qy 1186 SDRKDLVMGICICNNPPDDQHPFAVSPLEDDIINLMRLQINCESEULRLLYOFLVY 1245

Db 2065 EYLSDDSDQHHVLMKMMNTHIYDLYS--DKSTYNNV-----NFKKCINLLEEK 2112

Qy 1246 ONKANSVTLN--VSSNNGD-----IILANKVLQODNIKSMDDHNIHKK 1288

Db 2113 MHSBENKFFIKGFVEYNNADHINPEIPIKKKKKKKYKNNKNNSHKXKE 2162

RESULT 13

ID Q81CX0 PRELIMINARY; PRT, 2874 AA.

AC Q81CX0;

DT 01-MAR-2003 (TREMBLrel. 23. Created)

DT 01-MAR-2003 (TREMBLrel. 23. Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23. Last annotation update)

DE Hypothetical protein, expressed.

GN MAl6p1.48.

OS Plasmodium falciparum (isolate 3D7).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

NCBI Taxid=36329;

RP SEQUENCE FROM N.A.

RA Cherevach I., Davis P., Goodhead I., Stevens K., Mungall K.,

RA Berriman M., Pain A., Hall N., Actin R., Chillingworth C., Doggett J.,

RA Ormond D., Sanders M., Hayes K., Hall S., Quail M., Barrett B.,

RA Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.

DR EMBL; AL844505; CAD50319.1; --

KW Hypothetical protein.

SQ SEQUENCE 2874 AA; 334409 MW; 65541406B9D8BC CRC64;

Query Match 12.1%; Score 830; DB 5; Length 2874;

Best Local Similarity 23.6%; Pred. No. 1.1e-25;

Matches 323; Conservative 213; Mismatches 412; Indels 422; Gaps 56;

Qy 206 NSQNDSSRSALINIDKIDINKYSLGKIAPIDLASGRADPYSON-----KQOT 258

Db 33 NSRDNDPNDN-LNNNDPNINDNNNN-----DNPNNNDPNNDPNNDPNNN 83

Qy 259 DGANINRSIL--ALKECIKRAMDSD-----KNHLPFRDSELTKVLRDIFVGSKSIMTAN 310

Db 84 DNPININEQYNNFLFKEGDNYIDELVYTTNEPFIYAKEL--LDDKGLGDEHLPICS 141

Qy 311 ISPTISCEQTLNLRYSRYVKNFN-----KSTCINEEDPTNTERISILDSKSEWNA 364

Db 142 -----NKKFPELEKININPEDNDQYVSEFDFNRRRLNDEHFNNEKUR 189

Qy 365 SSIENVVLKSNHLSNNNNNNKINRKINDKE--RNHILKKSKSDKREGFTSIFGKYS 421

Db 190 YKNRHEFFYNN--NNNNNNNNYNNNERDVEYGRNL--NDSPD----- 231

Qy 422 SLINDIDKIKKKK--GLINYKSTLYNNTIKKNNNNNNNNNDNDNNNNNNNNNS 479

Db 223 SLANNDDNNNNRRSRFSIRANKFKFCHDNNMNLNNNNNNNNNNNNNNNNNNNNNDHNR 291

Qy 480 SSMVNMIMHMINNNINNNNNNNNNNN-----NNSHNNHLP 520

Db 292 NSYKKS--NYATLNDNNNNNNNNLNNNNNSGFKGLLMRDSDKTGSMTWRLGKSVSNH-- 347

Qy 521 QPNTVFTDTSDDSLDDNNCHLN--NNDKSLFLHKGLRDNIKLRSSCDIINAKKON 578

Db 348 -----NSMNHNNNNNNKLNNTNNMMSHNN-----KMSDNKRGHKTFFKNYNN 391

Qy 579 LHLARHSVSKLTWESYDPQKNDKTFPFKSNINKEEDTPKDIKESRVSNNNGAVLLG 638

Db 392 --DAKSNISSDM-----KNSISHEINVDN--NNYMHHLVNNNDTQGN----- 431

Oy		639	LANKTHADDIETKPEKEN-----HNDKKINNGVLTINIINSNVASINSMSNINSNM	688
Dd		432	-NKQNTY----TKDRIINIIMAGREDIYYTNEBHYTDG--LMTNYITMHPNN--ANIINNKN	481
Oy		689	NSNSITSYKSYNVSQOSISDVQIRYVENM-----DTSKNNDNI---FP	727
Dd		482	NNNNKHANNFGNN-----OLPYIIQMNNNESGAGAGAPNDIVADHKRATVEFNPSPF	535
Oy		728	DAISCDNN-----MY-----PHTTN-----	743
Dd		536	VSVAKNGNIQNSSNNYYKGVNNVTYKSGSEMTNSTSYGLSPININMMDMKKRYKKFESN	595
Oy		744	-----NNNNNNNNNNNNNNNNI-DVENKNNSDGTSTSMKLAYANSH---	781
Dd		596	NNNNSSMNVANNNVANNVANNVANNVANNVANNVANNVANNVANNVANSAAHTNNHHHL	655S
Oy		782	-----NLFPDPNNK-----NTSNIQNIINTKKNQDGAVN	810
Dd		656	NNAGVHESTRILEEKEFKOSIQNEYEPNNLYMKNNEMNNMGNSINLIRDNNNNNNNNNFN	715S
Oy		811	YSNMFCHTYNLDKDYLLIDLNKEQDKONHGCDDNI-----IQNRNDEKKKTTPYNN	864
Dd		716	YNNPF-NYNOHQYGQKSLINVNTYNNIINTHNANNTPYNNVNVTNNNDVVEKNGYYKDS	774
Oy		865	NNIYIVANNNGNNNSPRMKYGLCGSHTSIDMKQNNEMKNNEMKD--NEMKONHISKNNNN	922
Dd		775	SNNKNNNNNNIGNNT-----TYVENMENNILAKKSEKSNTPSHHSFDNAKNOTTN	823
Oy		923	SSSSSSSSNNNI-----YNNINDDTFOXDYCHANDFTTIRRKONTINSNIYQNDITI-Y	976
Dd		824	FTGSHYNNKMLRTNEKNSKNEKSEVIEDDWS-----LIRKK--SSIIDNKKASDIDLHM	875
Oy		977	TYSNLNDYMSNTLIHFPEKYTYPTJSTN-----EDIYKMEGGKHIRLDGDXY	102
Dd		876	KMSQGDFRNKSKTLEISEKKQOVTIKDNKFEPAKKNKESKDKCKEGCVNNKKDEKN	935
Oy		1026	DDNDNNNNVDNNNNK-----NVDNNVNDNNVNDNNVDNNDKNN	106
Dd		936	NNNNNDPDPOKNNNNNSNYNTGTPLTNTAMTGVKRGNTSNNNNNNNNNNNNNNNNNNNNN	995
Oy		1062	VDNNNVDN-----DDDDVDPHN-----IKCFNNNEYLSYF-----	109
Dd		996	NNNNNNKSNRMYGATSNITKEHPDRNNNYNKDKIMANNMSYLVKKNONE--SVFTNGSSN	105
Oy		1092	-----OKNVDITI-----NCLNSISL-----ISSMYDTRKILN	112
Dd		1054	INSVNNMFEKENENMVEKKKKKVVELNTHGEKNDNDLNSKDGNSPSKJNSQQNNNNNNNN	111
Oy		1121	NILLSKYKA-EKDONVIKITYINE-DIKONSLEBIERTQAOSIYEKRVYLTLGLLFFKQV	117
Dd		1114	NNKQKKKKKKKEKEDONKLTKNVEDNDPVKKLYALDNNMASSELKOXO-----ILLNDE	116
Oy		1178	DTQIINETSLRLKDYLMCHI CANNPPDQGFHYAYSRLKEDIINILMLERQWCSESENRL	123
Dd		1167	KTNNNNHNNSSHASYV-HENVYFGNQIINNYNNKQOONNIITYEL-----DHANNNDL	121
Oy		1238	YOFUVEYOKXSANSYVLWVSSNGDILINKCLVODNIKMSMOBNHIK	1287
Dd		1219	MOKNIMLMNDNIPODSLQIKKNNNNSVENYALNALATHYAKVMEMWNISK	1268
 RESULT 14 081519 PRELIMINARY; PRT: 2432 AA. AC 081519; DT 01-MAR-2003 (TREMBrel. 23, Created) DT 01-MAR-2003 (TREMBrel. 23, Last sequence update) DT 01-MAR-2003 (TREMBrel. 23, Last annotation update) DE Hypothetical protein. GN PF1085M. OS Plasmodium falciparum (isolate 3D7). OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.				

OX NCBI_TaxId=36329;
RN [1]
RP
RC
RC- STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Anguilo S.,
RA Perlea W., Allen J., Selengut J., Hart D., Mather M., Vaidya A.B.,
RA Martin D.M.A., Fairclamb A.H., Frumholz M.J., Roos D.S., Ralph S.A.,
RA McPhadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum".
RT Nature 419:498-511(2002).
DR EMBL; AE014847; AAN36303.1; -.
KW Hypothetical protein.
SQ SEQUENCE 2432 AA; 284064 MW; EF0989EB11BD2A33 CRC64;

[illegible]


```
QY 655 -----NDNKI-----NNGVINIINNSNVNS-----INNSNMSI-----NNSNMSNSLYKS 696
DB 1255 NIFSNVAVTNDTDLJACNNEORNNI SNEHINKNVYNNLNVASINKEIINNYFINSTODYNT 1314
QY 697 NYSNOSISDQIRYVEMEDTSNKNNDNI FPDALSCDN-NMYPIVITNNNNNNNNNNNNNNNN 755
DB 1315 NYDKNGCGLTLCIKNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN 1374
QY 756 NNIDVENYNRDGTN-----NSMKLYAVN-----SHNLFOPDNKN 791
DB 1375 KDIILKNKNNINNNANIFINISLKEKLLNDISKNKKKKKCKVIRIKKPKKVPKPKNN 1434
QY 792 TSN-----IQNTNKNKNDG----- 807
DB 1435 INNRKNGSKYKCDTRVKNKINNGPKPKKNTYSSCINNYKCALISDSTIYIKNKEFVI 1494
QY 808 -----NVSVMNFCBYNLN----- 821
DB 1495 GNYPTNGSVYITNGEANTTNNGKKNDHRKKKKKYNNENNNNYCNNSNNSNNNNNNNS 1554
QY 822 ----DKNYLIDLNKKEQOK-----NIHG----- 841
DB 1555 HFQGNKNNRGGDNDDEQKNNKRNIGKNGDDEEDYDGEDDNNNEENDHNNKNIKRN 1614
QY 842 -----CDNNIIONRNDPEKKKKTNF-----YNNNNIVYNNNNNGN 877
DB 1615 KKSNAKAYNNNDNENKCNKKOTFHLPKKHYEYTMNNNNYNNKNEYNKDEELKNKYEDNQ 1674
QY 878 NSPRKAYGLCSHTSIDNMK-----NNEMKNEMKDNEMKDNHISKNNSSS 925
DB 1675 DNGDILKHSCTNGNDVGVSKYKHNDHSGDNNNDINNSDSNNNNNNNNNNNNNNNN 1734
QY 926 SSSSNNNTYNNND-----DQTO-----NDYCHNDFTTRKANT 962
DB 1735 NNNNNNNNNNNNNNNCTNCGATVDLMLNLSNLKESYLTNYLSHNIN--NIPIINN 1792
QY 963 NINSNIYONDIIYITNSLN--DYMSNTLHFEKRYTPTLSTNE--DIYNKEM----- 1012
DB 1793 NHKKNFLANNENVOFPNNNYKDYVNN--KNTLEPNNTYINHNINELNFKKMLIANNLY 1850
QY 1013 -----EGHRIID-----DQKDYDNDN 1031
DB 1851 INYVGINNDRYMNGINNSQMSYNNHFRNDISNVLKGVYINKESHYKNNKNNKNGSNN 1910
QY 1032 NVDDNNKNN-----VDNVDDNNVDDNNVDDNDK--NNVDDNNVD----- 1068
DB 1911 NNNNNNNNNNDEDDDDNNNNNDNDGDNDNDNNNEENNNDRIIPFGHNNNNVSMKN 1970
QY 1069 NDDDDVDPHNTKFPNNNEELSYFOKQVDTIINNCLNSLDISMYDTEILLNNILSKYK 1128
DB 1971 NDGKNEDNNNNNNNNNN-----NNNNNNNNNNNSNDNNNNKNNNNIIPFLYNNK 2022
QY 1129 AEKDNVYIKKYINEDIKNMSLEBIDTAOSIYEKRYVLLTKLLLFKQVDTQINNETSDL 1188
DB 2023 FCMNNNNANNFNNNGI-----INNFTNN-- 2044
QY 1189 RKDLYMCHICNNP-----DDQFH-FYAVSRLEKDIINLIMRQIWCSEBENRLXYQ 1239
DB 2045 ----GMAH--NENPMNNINSFNNMMHLEFKYGIQYKDLIPNMMEPTA-----LNLGE 2093
QY 1240 FLVVEKONKSAN-----SVLLANSNNNGDI-----LANKKLVODNIK 1277
DB 2094 RRSKYSKSSYNNKKDLMKRRYELQKAILLVNDENLKEVYIDEIIRNSTILLPQKIRG--R 2151
QY 1278 NSMDHNN-IH 1286
DB 2152 NTLDCNHPH 2161
```

```
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Asparagine-rich protein, putative.
GN Pf1530W.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyte S.,
RA Chan M.-S., Nene V., Shallow S.J., Sun B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Valdivia A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McPadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrall B.,
RT "Genome sequence of the human malaria parasite Plasmodium
falciparum."
RL Nature 419:498-511(2002).
DR EMBL; AEO14849; AAN36392.1;
SQ SEQUENCE 1824 AA; 212658 MW; 6CECDE0EA6960084E CRC64;

Query Match 12.1%; Score 826.5; DB 5; Length 1824;
Best local Similarity 23.2%; Pred. No. 1e-25;
Matches 376; Conservative 244; Mismatches 450; Indels 549; Gaps 76;

8 VVRKPLSELEKKKXSDITVKNCTL-YI-----DEPRYKVM----- 46
DB 191 ILKNPLMBQNDKSN-----NVCTLIYIKKKTSLSDNHSRIKNEKEHINQALV 243
QY 47 ----TKYIERHEFYDKVPDDTVNDFVYENTIRPLIIDLYENGCVSCFAYGOTSGKTY 103
DB 244 RSKNPVASESLKEIKEQFEEIKQEDLYVDEIGLLILSPDNC----- 287
QY 104 TMLGSO-----PYGOSDPFGIFOYAGDI-----FTFLNIY--DKONTG----- 141
DB 288 TLVDQIKDYVLSKGD--IEMFLRGEIEKIKONNNNNLNYLLIKMONEKMDLFPNNNDM 344
QY 142 ----IFISFEYELGCKVLDLQKRVNVALENG--KKEV--VKDLK--ILRLVTKREL 190
DB 345 LNDIKISFTDF-----EKM-----NGEYKEFEHIEDKRRKIYIILCKQI 390
QY 191 ILKMDGVLLRKIGVNSQNDSSRSHAILNIDLKINKNTSLGKIAFIDLAGSERGADTV 250
DB 391 IYILEE--LYKNVYKNTIKOE-----KYINKLYNNKFFYI----- 424
QY 251 SONKQOTDGCANIRSLALKECTIRAMSDSKNHIPRODSELTAKVLRIPIFGKSKIMIAN 310
DB 425 ----LNKANNNSFFNNNTIV-----EKNNH-----IVNVEMLKRETEQVES 461
QY 311 ISPTISCEQTLNLRYSRYKQNFKNKSTCINEEDPTNTERIS--ILDSKSEMN--ASSI 367
DB 462 VLENV--LENVLDEBQSFKEKIKINIMRNFLEDKYVEKIDIHSDIFENNHYLSND 519
QY 520 TNVSLSEN--ASKKTQINLNLNKQIS--NCFSKONNNNNETQFSLSTCLTESSFMD-K 574
DB 428 KIKKQKKKGLINYSTLYNDNTINKKNNNNNNNNND----- 464
QY 575 KKEKENSCK-----ETNITNEQNGEDYSNDILNNDMMNVNVCDSFTMKIHYNDQFNH 629
DB 465 ----NDNNNDNNNNNN--NNDSSM--VNNMIMHMINNNINNNINNNNNNNNNNS 514
QY 630 YPIITNNNNNSNNNNNNNNNNNNNNIYHISNNNNNISNNNNNNNNNNNNNNNNNNNNNN 689
QY 515 HNNHLPQPNVAFPTDTSDFSSLDNNCHLN-----NNDKSIPLHKKNLNDNITKLX 563
```



```

Db 690 QFNTHNNNNLIINOYDQYNI-PLNNHNLNQLTTOGNCNNNELS-----KQLSNQFNQ 743
Qy 564 NRSSCDNIMNKKKNNLILARHSVGSKLTMFSYDP-----QKNK----- 601
Db 744 VKSLNNIVGSANN---LVNEYFSDMYIYSWVKSACQYPRELPSAERKLSEYIV 800
Qy 602 -----DNTEP-----KSNIMKEDNTP---KDLIYE 624
Db 801 AKEQMLYLRVWGYDEIDNIYFHPKSGSHIKIFKFSMCMNNFLKAYKNTDCKEMDFR 860
Qy 625 SRNV-----SNNGVNLGLNKK-----THDIST- 649
Db 861 PFNIPRSSISVRLRIERAVPRSSAAEFKKMKLNKYODLFLPKCNFVCEEDMOTIE 920
Qy 650 ---KDEHNDKINNGYININNSNVNSINNSNNSINNSNNSIYKSYNSQISD 706
Db 921 NMEGLDHYDTYNKQVILQNKNDITNKLNEQSLSNMGMANINIHDEYPPNN--- 976
Qy 707 VOIRYVEMDTNKNNDNIFEDALSCNNMYPNITNNNNNNNNNNNNNNNNNNNNNNNN 766
Db 977 ---INNINNKNNNIILLD-----EHHHHNNNNNNNNNNNNNAI----- 1014
Qy 767 DGTNSKLYAYNSHNLFPDNN-----KNTSNIQNTKKNQOGVNYSNFCHYN 819
Db 1015 ---VNSLSIPYAGINSYAPGVNHGFLGSNNNAVMMNTITNKTNDILN-----NLCSS 1067
Qy 820 LNDKNYLIDLNNK-----EQOKNTHGC----- 842
Db 1068 TGAKNYKMDINNLRFVGTKIDNNTLNINTDHFNTDVMKNHNTLTITTTTTMTKNENDI 1127
Qy 843 ---DNNTIIONRDFEKKKKTFYNNNNNIYI-----VNNNGNNSPRMKYGLCSHTSIDN 895
Db 1128 MNYNNLHNMTMFASTRSSSRNNNNNNININSINIVNNNNNNNN-----VNNNNNN 1177
Qy 896 MKNEMKONE--MDO-----NEKMDHAKSNSSSSSSNNNIYNNI-- 938
Db 1178 INNNGSNNNYLLKKSILNRGSLSLNLFNTNYYNNNNNDNNNNNNNNNNNNNNNNNN 1237
Qy 939 --DDDTQNDYCHNDTFTIRKKNNTNINSNIYONDII-----YTINSLN--DYMSNT 988
Db 1238 SSASVNLNANLNNINSNTI--NNNTMNSTNIMNNNLSNNIILSNNNVNNNTKDFLAN 1295
Qy 989 LHFKEKYTPPLSTNED-----IYKMEGHILRDQKYDDNDNNNVN-- 1036
Db 1296 -FNLKNTHTASLTSNGNNNLNEQFISYNSNTE---NINNNNSNNNSNNNSNNNN 1350
Qy 1037 ---NKNVND--NVDNNVNDNDNDKNDVNDNNVNDN-----DDDDVDFFN 1078
Db 1351 SNIMNSNNNSNNNSNNNSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1410
Qy 1079 IKPNNNNEYL-SYFOKRVDTIINNCLNSLDISSMYDTKEILNNIILSKYKAEKDVYIK 1137
Db 1411 INNHGHNMAHADKRV-AVYNFLNNIDINS---KKE--NLLLANLYK-----NS 1456
Qy 1138 YINEDIQMSLEEDKTAQSIYERK-----KYLTLKLLLFKKNDVTOINNETSD 1187
Db 1457 YLANN--TNSGNTFSNGQGVNOLKCLNTYNNKLNEBELBELNEK-----YNNNE--D 1507
Qy 1188 LRKDLVWCHICNNPDQHFHFAYSLKEDIINILMRQWCESE-----N 1233
Db 1508 LKDDL-----KQLINIDFINDIDMEYEEGHSYNNETFLRTSN 1544
Qy 1234 LRLLYOLVYE-YONKSANSVLNVS--SNNGDILANK-----KLVONIKSMOHN 1283
Db 1545 SSLQRTTYPKNPSFVENTLLTYVSTSNNGNINIKREYQTYDKLPYNNVMNCDEQN 1603

```

RESULT 16
 Q81EE4 PRELIMINARY; PRT; 2672 AA.
 AC Q81EE4;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

```

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Batteil B.,
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL044509; CAD52316.1; -
KW Hypothetical protein.
SQ
SEQUENCE 2672 AA; 310742 MW; 8929A8A4AC08464 CRC64;

Query Match 12.1%; Score 826.5; DB 5; Length 2672;
Best Local Similarity 22.5%; Pred. No. 1.5e-25;
Matches 368; Conservative 215; Mismatches 410; Indels 641; Gaps 72;

Qy 158 LQKRWYALENKKKEVYVYDKILRYL-----TKRELTKMIDGLRL 201
Db 1 MKERKYSKVESKSLVENDNFCNNYRNPFYMENDKKNMKEKEE-----R 49
Qy 202 KIGVNSQDESSRSHAILND---LKDINK--NTSLGIAPIDLASER---GADTV 250
Db 50 KEIFTSRNDG-----LVVDYNNKMMNDINKYENTILN-----SNNSCNNPFPYDQI 97
Qy 251 SQNKQOTQGANINRSLLAKECIRAMDSKNHIPPDSLETYVLADIYVGSKSIWIAN 310
Db 98 IYNNVYLSNGP-VNNS-----VNSNTNTNNMM-----NNNNMNSN 131
Qy 311 ISPTISCEBQTLNTLAKSSRVKPKKSTCINE-----EDQNTERTISLDS 357
Db 132 M-----V-KNNNNNNNNIYN---SNMKNMMNDMMNDMMNDMMNDMMNDMMNDMM 177
Qy 358 KGSNNASSIENVIKSNHLSNN--NNKINKGINDKIERNNIIKNSFDK----- 408
Db 178 MNDNMNDMMNDMMNDMMNSNNIMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 237
Qy 409 PRGGFTSTGKY-----SSLND-----IDKIKKKKKGLIN-----YYS 442
Db 238 NKEPFIOTQNMVLYKDNANNLNSCYVNTSMCDSCSNNSNVSNMVDKITVDHSCMSNNS 297
Qy 443 TLVNDTP--INKGNNNNNNNNND-----NDNNNDNNNNNNNN----- 477
Db 298 NDHNNNSNCYVYNNNDYNDYNDYNDYNDYNDYNDYNDYNNYNNNTKGQEGQKOG 357
Qy 478 -----DSSMVAANNINH-----MINNNINNTNIVNV 503
Db 358 KIGNGKLVMMDNMFPNINEHNNIYRERIDSLNTYNTSNYKNNYVNIINNDERNNNNNN 417
Qy 504 NN-----NNNNNNN 512
Db 418 NNANYKMMENPEKTYTNKKYAEENILKDSNDIHVVNDVSLFNTCDIVSNPYNNNNNN 477
Qy 513 NSNNHNLPOPNYAFDTDSPESSLD----- 536
Db 478 CGSSNNIHSNPLVENTSSFCEENHTSIYPPKKNKNNNNNNNNNNNNNNNNNNNNNN 537
Qy 537 -----DMNGCHL--NNNDSIFLHKNLRDNIKLNRSNCDN-----INMKKKK--NLH 580
Db 538 FPEGTHINCDLSNNNNYAVLTNKNNSNLNBSINVSCTCANNNDPFHANNNNNNNNYNIQ 597
Qy 581 LARHSVGSKLTMFSDYDQKNDKTFPKSNINKKEDNTPPDILYESHVS-----NNN 632
Db 598 IKDINVINIKLENYFENVLNNNNNT--KEYINNTYMKHIDIDTQHMNVHLNNSIKSIND 655
Qy 633 GNV-----LLGKNTKTHDIST--KDEHNDKINNGV-----NINNSNV--NS 674
Db 656 NNLEHINPMMETQPTPHI FNNIILNNSNNSNNSFSVSNIKETIRNNITDSSVIRNL 715
Qy 675 INNSNM-----NSINNSNNSNSIYKSYNSQISDV----- 707

```



```

Db 716 LANNQMHKLVNNNNNNYINDDIKMGDGGVAVLQNNPKMTSVTAVTSYSLGNV 775
Qy 708 -----QIRYNEKD-----TSKKNNDNIPFDALSCDNNMYPIITNNNNNNN 749
Db 776 LENIKRPSQYNNMNSYLGNNVKYINKDHELSF-----NKEDYICNNNNNNNAENTK 828
Qy 750 -----NNNNNNNNIIVENYNNRD--GNNNGMKLYAVNSHLFQPD--NNQNTSN----I 795
Db 829 NNNINSNDNNNDNNNDNNNDNNNDGNNN-----YNNHHLKHKDKINNPNFNSWIRKL 882
Qy 796 QNITNNKKNODGN-----VYNSWNEC-----HYNLNDKNYLIDLNNKEQDKNIHICD 843
Db 883 KSPGASNDENDENELPINAHNFYINFLNGHFNNKNDINIRINSKSGNSISE-D 941
Qy 844 NNIIQNRDPEKCKKKTNNYNNNNIYIVNNNGNNSPPMKGLGCSHTSIDMKKNENKN 903
Db 942 NDM--NNNNINS-----NNNNINSNNNNNNNNND--NNNNNNNNNNNNNNNNNNNI 990
Qy 904 NEMKONEMKDNHISNNNSSSSSNNNNIYNNIND-----DTEQN----- 945
Db 991 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1050
Qy 946 -----DYCHNDYT-----FTIR----- 958
Db 1051 LPKIKGVAFDQSGRRVWASWGNGNQKQYFVKKFGQAQAKYLAIYARIGAVNCLQKP 1110
Qy 959 -----KNNNTINSNY-----QNDIITYINSNDY 984
Db 1111 HKPWSKELSEKKKLLNMEGKKNELGNNECNNNNNNYISKNGEDD-----NIESY 1165
Qy 985 MSNTLHFKEKYTYPTLSTNEDIYKMEGKHRLDOD-----KYD- 1026
Db 1166 NDDEICEEENY--DIDHND-----EHYDIDDGDGNDGNDGNDNDDEHDI 1215
Qy 1027 ---DNDNNVNNNNKNNVNDNV--NNNVNDNVNDNNKNNVNDNNVNDNDVDFHNI--- 1079
Db 1216 TNNNDNNNDNNNDNNNDNNNDNNNNNNNNNNNNNDNNNDIKMDQSGDKEYGILNN 1275
Qy 1080 -----KPFNNNEYSTFO-----KNVD-----TINNCNSLDIS 1109
Db 1276 ESKNVGNKIGNIGNIGNIGNGNVGNKIGNKVEKIGNPIDKDEYMKSIITNNDDTCKOD 1335
Qy 1110 SMYDPTKELIANNILSKYKAEKDNVYIKYINEDIKNM--SLEEDTKAOSIYEKAYLL 1166
Db 1336 IANETTKIIPNNITDNNHIVDNCVCLMASNNLENIPEDKKGEDVTHISIDEQKD--Q 1393
Qy 1167 TKLLLLFFKN-----VDQIINNETSLRKDLVNCCHICNNNPDOQFHFYAYSRLK 1216
Db 1394 SNKFLRERKKNKMHNNYDNFYONSFKKLDYNEEI-----NKSSDE-----EK 1437
Qy 1217 DIINILMROJWCESENRLLYQFLVVEYQKNSAVLANSNNGDITLANKVLQ--D 1274
Db 1438 EVSQTNSLR--YKCEEN-----DQYDKQERKILQ--DTNCEIICINLESTKOD 1486
Qy 1275 NIKNSMDNNHTRK 1288
Db 1487 NIKDL-----SLHKE 1496

```

```

RC STRAIN=AX4;
RA Gloecheer G., Eichinger L., Szefranski K., Pachepat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Turguel B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.,
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC115682; AAL92689.1; -.
DR InterPro; IPR000629; DEAD box.
DR PROSITE; PS00039; DEAD_ATF_HELICASE; 1.
KW Hypothetical protein.
SQ
SEQUENCE 901 AA; 10285 MW; C3B695494117057B CRC64;
Query Match 12.0%; Score 824; DB 5; Length 901;
Best Local Similarity 26.5%; Pred. No. 6.8e-26;
Matches 278; Conservative 139; Mismatches 289; Indels 344; Gaps 40;
Qy 417 FGKYSLSNDIDKIKKKKKGLIYKSLYNDNTINKKNNNNNNNDNNNDNNNNNN 476
Db 2 FSKFS-----IKENPKKYKSNSSKGFNNNEINSYNNNNNNNNNNNNNNNDSDPEINN 54
Qy 477 NDSSSWNNMIMHNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 520
Db 55 NDS-----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 101
Qy 521 -----QPNVAFDTSDFSLLDMCHLNNNDKSIPLHKQLRDNIKLKNSR 566
Db 102 SGSDKFSIKPKKIYTPKRNKSTGSDIIDFNCCNNNNNNKSNDSNNNNYNSN--SNN 160
Qy 567 SCDNIMKKKNNHLLAHSHVSKLTMFSTYPOQKNDTPFKSNINKEEDTTPDILYESR 626
Db 161 SCNNNTN-----GNERKMGYGYGNNSNNNNSS-----SR 192
Qy 627 NVSNMGNVYLGLNKHQTHHDISTDEHNNKLNNGVYIINNSVNSINNSNNNSINNS 686
Db 193 DY-----NSNNNSSPDYNNNN--NN-----NNNNNNNNNNNSPRDY 231
Qy 687 NNSNSIYKSNYSNOSISDVOIRYVEMDTSKNNNDIFEDALSCDNNYPIITNNNN 746
Db 232 N-NRDNVNNNNNNNNNSPRD-----SNSNNRRSSRDFSNNSSPDYNNSRN 279
Qy 747 NNN--NNNNNNNIDVENTNNRDGTNNSMKLYAYNSHNLFPDPNNKGTSTNQNITNNK 803
Db 280 SNNSEFLNNNNNNNN--NSNNSNDNSPRDYVSSNSSPDYNNNDTF--SKR 329
Qy 804 NODN-----VYNSWNEFYNLNDKNYLIDLNNKEQKD 836
Db 330 NNVGNSILEIKSNPTSPRIPIWPPRSSPRIYV-----KLPIVQSLVDLEIMID 381
Qy 837 KNHGCNNITIQNRNDEKKKKKTNNFYNNNNIIVYNNMGNNSPFRMYGLGSHTSIDNM 896
Db 382 DNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 429
Qy 897 KNNEMKNNEMKNNKNNNNSSSSSSNNNNIYNNINDDTFQNDYCHNDYFTI 956
Db 430 NNNNNNTNN--SFNSDSNNYNNNSNNSNSNSNSNSNSNNNNNNNNNNNNNNNN 488
Qy 957 RKNNTNINSIYONDII-----YTINSINDYMSNT----- 988
Db 489 NNSNNSKSNNSNNFNIKLEGRKKSINEISDNLTRRIEIDSGMAHEFHEPLEKILFA 548
Qy 989 -----LHFEKXYTY-----PTLSTNEDI----- 1007
Db 549 GQRPVECDYFLKRYKQFQYVLDISMKISINTRODIGEDDDEINSVPFKCLCYE 608
Qy 1008 YNKMEGKHTR--LDODKTDNDNNVNDNNKNNVNDNNVNDNNVNDNNKNDNN 1065
Db 609 YDEENNNMENTGFIDEADKTYNRSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 668
Qy 1066 NVNDNDVDFPHNIKNNNNEYLSYFQKNVDTIINNC--LNSLSDISSYDPTKEILNNILL 1124
Db 669 NSNNNNNN--NFDYKXYFKN--YFQHLISKHILQIKAKCIQOTYVNDAA--NIPL 719
Qy 1125 SKYKAE-----KONVYIKYINEDIKNM--SLEED-----KTAQSIYERKKVLLTKLL 1171

```



```

Db      720 GSKKCHLSESEILIKNEFTKIKLHHHDEIDIDICPRCKIKKSIQKRIYFESK--- 776
Qy      1172 LFKK-----NVDQIN-----NETSDLRKDIWMCH 1196
Db      777 --KQKQXDDNNKSPVSNKVAATISHNPSNNIDKNNNNKNNKNTNNLEND----- 829
Qy      1197 ICNNPPDOHFYAVSRLEKDIINLIMROIWCESENIILLYOFUYEYONKANSVTLN 1256
Db      830 ---NNDSNNNNYLENEEPEIY-----FQDIKPPNNNNNN---N 863
Qy      1257 VSSNNGDIILNKLVODN---IKNSMDHN 1284
Db      864 NNNNNND---NNNDNNNDNNNDNNNDNN 889

RESULT 18
081107
ID      081107      PRELIMINARY;      PRT; 1483 AA.
AC      081107;
DT      01-MAR-2003 (Tremblrel. 23, Created)
DT      01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT      01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE      Asparagine-rich antigen.
GN      pfl_0111.
OS      Plasmodium falciparum (isolate 3D7).
OC      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX      NCBI_TaxID=36329;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=3D7;
RC      MEDLINE=22255705; PubMed=12368864;
RA      Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA      Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA      Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA      Chan M.-S., Nene V., Shallow S.D., Suh B., Peterson J., Angiuoli S.,
RA      Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA      Martin D.M.A., Fairhead A.H., Fraumholz M.J., Roos D.S., Ralph S.A.,
RA      McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA      Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA      Fraser C.M., Batteil B.;
RT      "genome sequence of the human malaria parasite Plasmodium
RT      falciparum."
RL      Nature 419:498-511 (2002).
DR      EMBL: AE014837; AAN35697.1;
SQ      SEQUENCE 1483 AA; 173923 MW; F93DB19A2A34D5F1 CRC64;

Query Match      12.0%; Score 819.5; DB 5; Length 1483;
Best Local Similarity 25.5%; Pred. No. 1.6e-25;
Matches 319; Conservative 160; Mismatches 333; Indels 439; Gaps 57;

Qy      189 ELLIKIMDGYLLK-----IGVNSQ-----DESSRHAIITND 222
Db      11 ELPLKNDVDVDEKDVITNAGININSENSNGVEIKEGDEKDHIOENDDETIVNEVD 70
Qy      223 LKD-----INRNTSLGKIAFIDLAGSERGADTVSQKQOTQDGANINRSLALKECIRA 276
Db      71 LKDDVTVENKKE-----EKSSNRWADWCECDAPLVDTYNNNDMDK-----KV 117
Qy      277 MDSDKNIIPRDESLTKVLRFIVGSKSITMINISPTISCEQTLTLTYSSRVKFN 336
Db      118 VDNND---FNNSDMMNNMKQNYMAPKN---DMNNYATCYEYIYNLHIDI-----T 162
Qy      337 KSTCINEEDTTERISILSKG-----SEMANS-----SIT--- 368
Db      163 KEELIYLFDPYRKRIINILNKKGRTAAAYEPDIEIMDSLEINGMIYSGNNSFSGHT 222
Qy      369 -NVVKSNNHLLSNNNNNKIRNGKINDIERNNIKNSKFPKPRGFTSTFGKYSLLNDID 427
Db      223 ISVILNDKKNKFNHVNQKN--KFKKMTNNNN--NNNFSNIGK-----NNNN 267
Qy      428 KIKKKKKGLINKSTLYNDNTINKGNNNNNNNNNDNNNDNNNDNNNDSSMVNN-- 485

```

```

Db      268 NIMNN-----YGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 315
Qy      486 -----MINHINNINNINNNIVNN-----NNNNNNNNNNNNNNNNNNNNNNNNNN 524
Db      316 FPDNNKMMKMDMKVGVNTINNNNNQNFYNNIKSYLNNNQVGSIVHNNH--QQNN 373
Qy      525 APTDTSFSLDDMN--CHLNNDKSIFLHKRLRDNIKLKNRSCDNIN-----KKKN 578
Db      374 MLDKNNNNNAFNKQNYLNNN-----HNKQSLADN--NNNNNNNNNNNNNNNNNNNNNN 428
Qy      579 LHLARHSVSKLTMFSYDPOKADNTPFKSNINIKMEDNTPKOLLYES-----RVNS 629
Db      429 LILKKSVPLOKSYVNN-----NIFGEA--KPVIDPEKILNSTTEKMDNTSN 480
Qy      630 NNNGVNLLGLNKVTHDIDSTKDNND-----NKNNGVININNSNV 672
Db      481 NNNNN-----NNNSDDAEKKNENIDMEKKKQOQEGTDKYKLSHLN---NTNNNNI 533
Qy      673 NSINNSNNNSINNSNNNSNS--IYKSNYSQGISDVOIRYVENMDTSKNNNDIFPDA 729
Db      534 NEFNKN-----KNTNNSMKVNFYKKNNTNRFY-----YNNFANNTKNNNN----- 579
Qy      730 ISCDNNMYEYITNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 789
Db      580 -----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 612
Qy      790 KNTSNIQNTNKNQODGAVNYSMFCYNLNDKNTLIDLNNEQDKNIGHCDNNIION 849
Db      613 KNNNNLESNNNSKQNLNLNHS-----GNTNNSNNVNDNN----- 649
Qy      850 RNDPEKKKTYTFYNNNNNIYVNNNNNNNSPRKXGICSHHSIDMKNNEMKNNEMKON 909
Db      650 -----KKNNNVNNKGF--KKNMDNNNITR-----DFSLLKNTTPDNKNKIN 689
Qy      910 -EKNDHIIKSN--SSSSSSNNNIYNN--IND-----DTQNDYCHND 951
Db      690 VSVIHHIYENNNKLYGNOBEDSTNNNNNNNNYNNENGTSGNMKGKFNFFETHGMD 749
Qy      952 NTFTRKKNNTNNS--NIYONDITITNSLNDVNSTLHFKKRYTP-----TIST 1003
Db      750 KKNVDNNNNNTPYQKELNIIKDD--KVNIIIMK--IKRENNTAIPNSGVNLS 804
Qy      1004 NEDIYV-----KEMEGKI-----RLDDODKYD-----DND 1029
Db      805 SSNFRDSSNNRGGQKNEKAVENDLNNDRDGLKNMDDEBYDINNNMSKNTERY 864
Qy      1030 NNNVNNNNKKNVNNVNNVNNVNDNDV----- 1055
Db      1056 --NNDKNNVNDNNVND-----DDDDVDFNR--IKNNNNNEYLSYFOK 1093
Qy      925 DENGDEENDNNNSNIIVNGETGENEKYKDKYHDKLDNTNESYKSDNTYESDVSD 984
Qy      1094 NVDTIINNCNSLIDISMTYDTEILNIIILSKYKAEKONIVIKKYNEIDIKMSLE---- 1149
Db      985 TNNITDNNNNNNNDITS-----TNLKSITLQITKR--YKKTDTDETENNVENNEA 1035
Qy      1150 -----EID-----KTAQSIYERKRVLLTKLLLFKKNVDT 1179
Db      1036 NGDSNANAASGSKIDILIKPLKKGENIWEAR---THLLVEQGENAVT 1082

RESULT 19
081154
ID      081154      PRELIMINARY;      PRT; 4466 AA.
AC      081154;
DT      01-MAR-2003 (Tremblrel. 23, Created)
DT      01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT      01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE      Hypothetical protein.
GN      pfl4_0170.
OS      Plasmodium falciparum (isolate 3D7).

```



```

Qy 1273 QDNIN-----SMDHNN-IHK 1287
Db 2302 NKNINNDNINILNKAMDNNQKFIHK 2927

RESULT 20
Q8IHV8 PRELIMINARY; PRF; 2275 AA.
ID Q8IHV8
AC Q8IHV8;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein.
GN PF11_0418.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OC NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RA MEDLINE=22255705; PubMed=1236864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.B., Bowman S., Paulsen I.T., James K.,
RA Easen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Aginoti S.,
RA Partea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairhead A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA Mccadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrett B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).
DR EMBL; AE014842; AAN36001.1; -.
KM Hypothetical protein.
SQ SEQUENCE 2275 AA; 267960 MW; FC804F5EB2AED1B7 CRC64;

Query Match 11.9%; Score 817.5; DB 5; Length 2275;
Best Local Similarity 24.0%; Pred. No. 2.9e-25;
Matches 386; Conservative 229; Mismatches 449; Indels 546; Gaps 81;

Qy 41 RYKVDNMTYIERHEFIYDKVFDVTDNFTVY-----ENTKPLIIDLYENGCVSC 91
Db 54 RYKLNCKRRKRRK-----DENLDDFEFYIKWKYDSDEMTWP-----FEN---LSP 97

Qy 92 FAYGQT-----GSGKTYTMGSGQPYGSDTPGIFQYACD 126
Db 98 NLKGQARKVALQLTEEIDNCNKNKNEKNEGENIEETVGD-----NINRYIEGN 148

Qy 127 IFFELNIYDKONTGIFISFEIYCGKLYD---LLQKRKVVALEN---GKKEVVKD--- 178
Db 149 I---INNEMNSQGNV-----EYNNKVFQOTRLVQNTDVLNPKAKMGIEIKRPDIDD 200

Qy 179 ----LKILFVLTKEBILKMGDVLRLKIGVNSQNDSSRSFALIN-----IDLKD 225
Db 201 DDEEIIIMELKEENN-KKEKNKLRNFSLFNMSNSRLSNEINSLYENGSHKMDVTY 259

Qy 226 INKTSIGKIAFIIDLASERGADTVSONKOTQTDGANI----- 263
Db 260 IGGKKTGNIKSIILNTRINNDDNNNNNNNNNNIYNNKIYNNKRYNNNNYCKYSTDLSDY 319

Qy 264 -NRSILALKECIRAMDSDKNHIFP-RDSELTLYLRDIFVGSKSIMIANISPTSCCEQT 321
Db 320 GNNNSVHLNDLHLHNDSEKNNLFLYDSD-----DLI---KQDSRIAN-----DIH 362

Qy 322 LATLYRSSRYKNFKNK--STCINEEDPTTERISILDSK--SEMASSIERNVVIAKSN- 375
Db 363 YNNLFISKSIKJLRNKYLESLN-----NNINNKIIFDESLKKNKGLSGPSMLCINNNE 418

Qy 376 ----HLISN--NNNNKINCKINDKIERNNILKNSFPDPRSEFTSPCKYS---- 421
Db 419 KSKKCKQKPHFVNTVVKSNKNENIIND-----FNKEKDESSQNDTSFL---YSTDR 468

```

```

Qy 422 SLNDI-----DKIKRKKGLINX-----STLYNDNTINK 452
Db 469 STDYITRREKIKPLQKESIINNEEDVQEHKKLLDKGNKNNDDSDTDNNDDK 528

Qy 453 KH-----NNNNNNNDNNNDNNNNNNNNNDSSSVNMMIMINNNIN 496
Db 529 NNNNDKNDNDKNDNDNDNNNDNNNNNNNNNNNNNNNNNNNNNNSGDNNNNNNNN 588

Qy 497 NNTLVNN-----NNNNNNNNNSH--NNHLP-----QPNVAF-----TDPS 530
Db 589 NGDNCNNNDNNNDNNNDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 648

Qy 531 DFSSLDMMNCHLN-----NNDKSI FLHKRL--RDNIKLKRSSCDNIMNKKNNLHARS 585
Db 649 NSNSYNSNSNSMFEYNNININIDEGNKIARNGDLKRIK--NVLNCK----- 697

Qy 586 VSGKLTMSYDPOKNDKNTFFKSNINKMEDNTPKDIYSR-----NVSMMNGVILGLNK 641
Db 698 -GTRF-LPSY-----NTCFKKQICNLKRNKKKLLIVSMGSKFNISQVNNAN-----NE 743

Qy 642 NTH-----HDSIKDENHNDKINNGVININNSNNV-----SINNSNKSIN 684
Db 744 NNFEEFKQCLONGNDQTKKRLQHEKKKSTKSKSINNNNNNYDYDHYFNKKRSEDD 803

Qy 685 NSNMNSNISTYKSNYSNOSISDVQIRYV-----NE-----MDSNKNNDNIFEDAI 732
Db 804 NSPEFPDSYFTMSANTIDQVSLQVGYLALPENENIMATDYLNTTTNNND----- 855

Qy 733 DNNMYPNITN--NNNNNNNNNNNNNNNNIDVENYNNRQGTNSMKLYAYNSHLPOPDNK 790
Db 856 -NKNDNINNEENKNDNINNDNNNDNNDDNNNDNNNDNNNDNNNDNNNDNNNDNNND 901

Qy 791 NTSNININMTKNNOGNY-----NYSNMFCHYNLDKRY 825
Db 902 N-----NDNNKNDNDKNTCKKGRKTNNIKKAILVNDNNINLGNYSNNE--YTDHVA- 954

Qy 826 LIDLNNKQOKDN-IGCD-NNIIONRDPFEKKKTTFYNNNNNIYVNNNNNNNSPRK 883
Db 955 ----NNSNINIMHNKSLSPFIDEDKIFKQSGNIIYK--LFCNKKSISNNSAYVN 1007

Qy 884 YGL-----CGSHTSIDMKKN-----EMKN--NEMKDNEMKN-----HI 916
Db 1008 RGISEYAEENIDKFNINNQKCNFYFLNKELONIYASEKYGKSKGNLVSCKINDYI 1067

Qy 917 KSNN--NNSSSSSSSNNNLYNNIND-----DTFQDYCHN--DNFT 955
Db 1068 YNNNSGNNSSNNNSNNNNNNNNNNIRKSYINKLTESEYDSYETLNNRYCKNTNKS 1127

Qy 956 IRRKNTNINSNIYONDIIYITINSINDYMSNTLHFEKITYPTLSTNEDIYK--EM 1012
Db 1128 KXINDENGTNPEINDIKY-----INDNNNSSTNSRSLIHDLSND--FNKIGDEL 1181

Qy 1013 EGGHRL--DDQKYDNDNNNVN--NKNVYDNNVNDNNVD-----NNDKNV 1062
Db 1182 KRQSQITKENQYAHVDDIDDEEENETDKN--NNTIRSIQSDIIYKILNDEKSKS 1239

Qy 1063 DNNNV-----DNDDDVDYF-----HNKIFNNNEYIS 1089
Db 1240 QONVHILKONINESEFKKKGENDVFFSSDNMKSNITYAKKNSSESFNKHLWK 1299

Qy 1090 YFQ-----KNVD-----TIINCLNSLD-ISSMYDDTKEILN----- 1120
Db 1300 KXOHLNMDEIKKRNKSKSVSDSTNNENHNNNNNNNNNNNNNNNNNNNNNNNNNN 1359

Qy 1121 ----NIIISKYAEKONVIAKYINEDIKNNLSLEIDKTAOSIYERKRVLLTKLLLFK 1174
Db 1360 KQDYIYNLENIYLSALKEN--RTYNSKDL-NISKE--NENANGHIRSNSTYDTSIILNR 1414

Qy 1175 KAVDQINNETSDLRKDLVMCHICNNNDPDQGHFAVYARLEEDINILMLQIWESNLT 1234
Db 1415 KIPEISGEKENYENLRNLRRTFPHDSIE-----YSTRTP----- 1449

Qy 1235 RLLOYFLVVEYONKSANSVLNVSSNNDIILNKKLVODNINKSMHNN 1284

```



```

Qy 212 SSRSHAILDLKIDIN---KNTSLGKIAFIDLASERGADTVSONKOTOTDGANINSL 268
Db 586 SNAINGLLNTLNTLNINSINTINGSLGLGAMNNNGHTNNNNNNNNHNNNGAIAIMANNGM 645
Qy 269 ALKAC-----IRADSDKHIIPRDSLTLYLADI----- 298
Db 646 GINMSMNISGLSTLNMNNLGYDIRSIGSTSNINLNINSMELNLKSGISINISNC 705
Qy 299 ---FVGSKSIIMANISPTISCCBOTLNTLAFSSRVKFKKSCICINEEDTNERISIL 355
Db 706 DENMIGLKNISTYNGELIG---NMINNYYANAI---NYNNY---VNSINNTSSNSFL 759
Qy 356 DSKGSEMASSIENVVIKSNHLNNSNNNNKINRGKINDKIERNNI---LKNKSPDKREGF 413
Db 760 NS-----VNYNTAYAHGNNPISIPGG---NNMSTNNYYSIQNK----- 795
Qy 414 TSTGKYS--LNDIDTKKN-----KKGLINYKST---LYNDNTINKK 453
Db 796 YSNFGKYSNLYNVPISQSIENEVSAGBEDYCDIDKSIDNLGSSKSPANNKLLGKI 855
Qy 454 HNNNNNN-----NNDNDNNNNNNNNNNNNSSSVNMMIMHNNNNNN 498
Db 856 NNNNNNNIISSSSTSHHINFKNTISNNNSNINNNNLINSTISVHPANNP---NMNDTNN 913
Qy 499 INVNNNNNN-----NNNNNSHNNHLPQPNVAFDTIS-----DPSLDDMNCILNN 544
Db 914 LNIKRNKNBESFISDLSTNKGELANNMGDHYGEMELNLYOMNMIKREKYEDEKINNNNN 973
Qy 545 NDKSIFLHKKRLDNILKRNKSCDNIANKK---KNLHLARHSVSKLTWPSYDPOKN 600
Db 974 MSKNNM---NLNNSSISNYYNLLNNMNSNNDLIKNNNTTYMKINMSSININNYTNT 1030
Qy 601 KDNFFPK--SNINMGEDNTPKDIIEGSRNV---SNMGNVTLGKNTKTHDSTKDNH- 654
Db 1031 VLNTDLKYSMLNONGINVTNSTYTSANSIDNNNNNNNNLNGNINDDHNGNSGKINHT 1090
Qy 655 ---NDNKN-----NGVINIINNSNVNS-----I 675
Db 1091 NPTIINDNKINVAHSHKNTSIDGMWINGPINTNSNKYSBEKDKANTEDKIGYKA 1150
Qy 676 NNSMNSINNSNNNSNISIYKSNVNSOSIDVOIRYNEKDTSKN-NDNIFPDALSCDN 734
Db 1151 NDNNNNSNNNRNNNGKVLNNKNSKNG-----NNKSKSKSKKNNNDPNKASGDN 1201
Qy 735 NMYF-----NITNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 789
Db 1202 LNEPDYKENNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1252
Qy 790 KATSIIONINTKKNODGAVYKSNFCHYMLIDKQYLIDLNNKEOKDKNTHGCDNIION 849
Db 1253 SNNNNNSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1303
Qy 850 RNDFEKKKTNFYNNNNNIVIVNNNGKNNSEPRMYGLCGSHTSIDNNKNNKNNKNNKDN 909
Db 1304 SNNNNNNHSTLISNDNTIYSNNDNEDNS-----NHTTNNNEQDDKKNNNK-- 1351
Qy 910 EMKDNHFK-----SNNNNSSSSSSNNNNIYN-NINDDDTFQK-----DYC----- 948
Db 1352 -TEGNSHKFSKQHGNNMTRKNDKKNFVNSNMLVNSNVEDVNNFNFIINQHRHCNKNYI 1410
Qy 949 ---HNDTFPIRKNNNT-----NINSINYONDI--TYTI 978
Db 1411 NFIKEDQEVDOKEHLASNISSHMLNEDNSNEIMDKTSYNNPYNNNANAYKSDNITKLYR- 1469
Qy 979 NSLNDYNSNTLHFEKERYTYPTLSTNEDIYKEM-EGKHILRDPDQDYKDNDNNNNVNNN 1037
Db 1470 ---GEKSSSQHLIKNNFSSNNNSNTEGELIIMEDMKNKPNNADILGN 1517
Qy 1038 KN-----NVDN-----NDNNNVNDNVNDKNAVND--NVDNDDDVDVDFHAKIKFNNNEY 1087
Db 1518 ENLIGIKLNNYNNPNVHNTERNELNNNTMTNKNVNSCNNMLHKQINNPN----- 1573

```

```

Qy 1088 LSYFOKNVDTIINNCLNSLDSIMYDTEKILNNILSKYKAEKDVYKKYINEDIKMS 1147
Db 1574 ---IINIKYVHNINISNL-----FSKN--MSNVVASENV-----NINNNS 1609
Qy 1148 LEEIDKTAQSIYKRR-----VLTLLKLL-----LEKAVDIOINNETSDLRK 1190
Db 1610 ---NTNNTANNPNFKKSKKANLNVFQOSLNNNLNMLNPNFYLNENNSFNKIMNAYNSIN 1667
Qy 1191 DLVWCHCNNPPDOQHFYAVSRLEKDIINLIMLRQIWCESENLRLLYQFLVVEYQKSA 1250
Db 1668 TSYFNTNDNNNNNTQHYTSSPIPEHPADKVTQK---SSEN----- 1707
Qy 1251 NSVLLNVSNNNGDIILLNKLYODNI-----KNSMDN 1283
Db 1708 ---NNNINNSSTWYKEX-EEPNLPKKNKVVQKN 1738

RESULT 23
Q8IDR0
ID Q8IDR0 PRELIMINARY; PRT; 3848 AA.
AC Q8IDR0;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN pf13_0235.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxId=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Barriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.,
RL Submitted (SEP-2002) to the EMBL/GenBank/DBS databases.
DR EMBL; AL844509; CAD52559.1; -.
KW Hypothetical protein.
SQ
SEQUENCE 3848 AA; 454964 MW; 57D2700BEA3F5A9A CRC64;

Query Match 11.9%; Score 813; DB 5; Length 3848;
Best Local Similarity 24.0%; Pred. No. 7, 2e-25;
Matches 330; Conservative 241; Mismatches 435; Indels 368; Gaps 66;

Qy 126 DIFPLNIYDK-----DNTKGFISFYELYGKLYDLLQKRMVALFMKKEVYVXDL 179
Db 1288 DILNY--IYENPFILDNK-----YELDOVYNOQMLKKTNTITIND- 1330
Qy 180 KILRVLTKEELIKMIDGYLLRKIGVNSQNDSSRSHAILNIDLKIDINRNTSLGKIAFID 239
Db 1331 KILHNLKLDNSTHGMW---IPNISYDNKMKGSSNNH---NNDLNNNNNNY----- 1375
Qy 240 LAGSERGADTVSONKOTOTDGANINSLALKECIRAMDSD-KNHIIPRDSL---TKYL 295
Db 1376 -----YNNYYLNNNNFENKRVLELLSNNNCFNRINQFCNKFEPNLMTHNDIA 1428
Qy 296 RDIYFGKSKI--MIANISPTISCCBOTLNTLAFSSRV-KYFKKSTCINEE 344
Db 1429 DDISTGSNEDVNDIYNIRKTDILKRLMYLSNNYNNNIEVDNYILHSGQHDHILNEL 1488
Qy 345 DDTNERISILDSKSEKN-----ASSIENVVIKSNHL-----SNNNNNKI----- 386
Db 1489 NNNTVYYNNIKKKKKNFNNQQLPRTGRNEDFIENNHHIYTPNNIYNNNNNNKKLADSEIM 1548
Qy 387 -----NKGKINDKI-----ERNNILKKSDEKREGETSTFGKYSGLNDID 427
Db 1549 ENVMOTCONNNYQFNTESNDMLKNIPOENNFLLPUSNNEK-----FGEIINY-DKN 1599
Qy 428 KIKKNNKKGLINYKSTLYNDNTINKNNNNNNNNNDNNNN-----NDNNNNNNNDSSS 481
Db 1600 NIEKN-----INNFNNMYN--VNNNINGNANNGNINGNANILIGNNINGNANNNINSN 1652
Qy 482 MYNNMIMHNNNNNN--NNINVNNNNNNNNNNNNSHNNHLPQPNVAFDTSDPSLDDMNC 540

```


D	b		1653	NINS--NNNSNNSNNNSNNNSNNNSNNNSNNNSNNI-----NSNNNS	1701
Oy			541	HLNNNDKSLFPLHKKNLRDNIKLNRSSCDNIIMKKKNLILARHSVSKLTMEFYDPQKI	600
D	b		1702	NIYNN-----MYMNYKNNVLSNNNYINNINNE-----QHSPMNAYTGYGDLSSN	1748
Oy			601	KDNTPEKSNL-----NKREDTPDI---LYESRNVSNNGN-----	634
D	b		1749	VNHLPYNNNIMDRYNISNNNIIQENIKGIYPHIYSKHVMVAETNPFIYDVNPKYT	1808
Oy			635	-----VLGLANKTTHDIS-TEDENH-NDKNJINGV--INTINNSNVNSINSMN	681
D	b		1809	SMDKPQVEGYIDREERYDINIILNDENNINKNIIINNINDMNYDYDNNSIHSNNNNFD	1868
Oy			682	SINNSNNNSISIKSYNSNQS-----ISDVQIRYVENWDTSKANNNDIFPDALSCDN	734
D	b		1869	YIKKIDNNNNINDRVDFSRSYEGSFPLITNGPYEENNTSSN-----ILCNO	1919
Oy			735	NMPNITNNNNNNNNNNNNNNNNNDIVENTNMEDGTNSKLYAYASHLLPORDNNKATSN	794
D	b		1920	DV--NLTNNNPIYESNISNNYQQLSDNNNEGNIIMNKNFVFNNNVL-----	1966
Oy			795	IONTNKKNNQGCVNVSMAFCBYNLNDKRYLID--LNKEQDKNIKIGCDNNIIORND	852
D	b		1967	INNCITNDNNVDNNKKMIDSIMIDGMIDSN-MDSMMGNMNNEMNNM--ENNIENNKM	20233
Oy			853	FEKKKTTFYNNNNIYLVNNMGNNNSPMKYGLCSHTSIDMKKNEMKNEMKDNEMK	912
D	b		2024	GNNIIEENMGNN---IENMGAGN-----IENMGAGNIIIDNNMIDNNIIDNNMI	20733
Oy			913	DNHKSNN--NNSSSSSSSNNNTYNNIN-----DDDFOND-----YCHNDFT	955
D	b		2074	DNNIIDNNLIIDNNMIDNNMMYDHINNFTINSFMIDMSFEEDPKXNMIGHNEIDI	21333
Oy			956	IRKRNNTNNS--NIYONDDIYITNSLDYNSNTLFHFKEKYTPTLSTN-----EDIYN	10099
D	b		2134	INTLEDLINNSENYKHN-----TSYLANYONCKRFELKANSVEIQOTIMDSNTTYN	21909
Oy			1010	K-----EMEKGHIRLDO-Q-DKYDD-----ND-----NNNVNNDN	10373
D	b		2191	KITEENTNSHILLEEKIKTKTDCKNDPEIHDISHLAILKKPNSEKKKIANNHHNNK	22506
Oy			1038	KNNVNDNVNNDNNVDN'-----VNNDKANVDNNDDDDDVDFAHKVPN--	10833
D	b		2251	KNNKKKKDDLHEHNNEKRLONCHNEFLDSNDYNS--DNNY--NKDVPVYINNIKNTYCN	23073
Oy			1084	--NNEYLSYQKANDTIINNCNSLPISSMYDPTKEILNILLISKYAEKD-NVIKTYIN	11464
D	b		2308	NNNDIIS--KNVD-----NSSKNDISNEKDN-----INMKLKMLKITNYKLG	23513
Oy			1141	EDIKMSLEIBDKTAQSIERKRYVLTKLILFKKAVDIOINNETSDLRKOLWMCHIGN	12006
D	b		2352	KKYHGICYPDPRNGWSFFVKDGVRYYKKPFSSPK-----YGULLA--KKKCIEWRLKTL	24033
Oy			1201	NPDDQHFHYASRLKDOIULIMLRQIWCESENLRLLYOFLVVEYONKSANSVLANVSN	12663
D	b		2404	NPS--HAYSFSLSKAKEFEVUL-----NDN-----YKDVGITLYDNKKDSNNNNNIXKN	24503
Oy			1261	-----NGDIILNKKL-----YODNIKMSDHNHHKK	1288
D	b		2451	GNQDDGGNRDILYVNAFLNYNDENEKCFDNDNDNVKNDNEKNTYEIKIIRK	2504
<hr/>					
RESULT 24					
OBT2M5 PRELIMITARY; PRT; 2472 AA.					
AC OBT2M5;					
DT 01-JUN-2002 (TREMBLrel. 21, Created)					
DT 01-JUN-2002 (TREMBLrel. 21, last sequence update)					
DT 01-OCT-2002 (TREMBLrel. 22, last annotation update)					
DE Hypothetical 278.5 kDa protein.					
OE Dictyostelium discoideum (Slime mold).					
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.					

OX NCBI_TaxID=44689;
RN [1]
RP
RC SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Glocckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Qaali M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium.";
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL, AC115593; AAI92295.1; -.
KW Hypothetical protein.
SQ
KW SEQUENCE 2472 AA; 278497 MW; 30CCCF157D4008A7 CRC64;

Query Match	11.8%	Score 810.5	DB 5	Length 2472
Best Local Similarity	22.3%	Pred. NO. 66-25		
Matches 302; Conservative	191	Mismatches	400	Indels 459
				Gaps 41

Qy 214 RSHAILINDLKDINKNTSLGKIAFDLAGSERGADYVSONKQCOTDGAINIRSLALKEC 273
: : : : :
Db 291 QSTSLTLPDSDEGNNF-----NNNSNNNNNNNNNNNNNNNNNNNNNREDDNLK-V 338

Qy 274 IRANDSDKHHIPRDS-ELTKVLRIPIGKSKSIMINISPTISCCEDTINTLRYSRNVK 332

Db 339 IEEIDDEDDDYLSDDSYSSRPESSVDLNTQTKFTTADIISSI-----PLSSSFS-- 390

QY 333 NFKNKTCEEDDINTERISILDSKSGEMASSIENVVIKSHLLSNNNNNKINRGKIN 392
Db 391 -----TTTFEENSNNTSINSNKIKPLEPI-NISQOMNNINONNNNSNNSSN 436

QY 393 DKIERNNILKNKSPDKPREGFTSTFG-KYSSLNDIDDKKANKKGLINYKSTLYNDNTIN 451

Db 437 SNTNNGGNTOPKKR-----TFSLPLSLMLNNNNNNVNGN-----NSNNNNNNNIS 482

[illegible]

Oy 512 NNSHHNHLQPNYAFTDTSDFSSLLDDMNCMLNNNDKSLFLHKKOLRDNIKLKRRSSCDNI 571
||::|:
543 NNNNNNN-----NNNNNNNNNN-----GTSSSSSSA 567

Oy 572 MNKKNNIHLARHSVSKLTFMSYDPÖKKNDNF-----FXSNINKMEDN-TPKDIL 622
 : : : : :
 : : : : :
Dh 568 TAASAGTLLSPRIMSPRLTTVOAPASTAASNPLVTMSPBROSANINSLAPGSSKILE 627

```

QY 623 Y-----ESRVSNNMANGVLLGLNKNKTHHDISTKDEHNHDKINGVINIINNSVN 673
      ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Dh 628 FHSMSKSKSKSKSDSDPDPTDGLSTIEFDYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 687

```

[illegible][illegible][illegible]

```
QY      803 NNQD-----GNVAVSMFGEHYNLNDKNYLIDLNKKEQKDKNIGCDNNIIIONEN----- 851
```

QY	052	-----DPEKKKKT-----	859
----	-----	--------------------	-----

QY 860 ----- 859

[illegible]

KM	ATP-binding; Coiled coil; Microtubules; Motor protein.	
SQ	SEQUENCE 1200 AA; 138263 MW; 7673746EFB604614 CRC64;	
	Query Match 11.8%; Score 810; DB 5; Length 1200;	
	Best Local Similarity 25.1%; Pred. No. 3.2e-25;	
	Matches 309; Conservative 201; Mismatches 370; Indels 352; Gaps 50;	
OY	3 SKIKVVRKRLPELEKKKKDSIIITVKNNLTLYIDEPD---YKVDTKYIERHEFIIVD 58	
DB	-256 SNVKAIVAIKIGEESE-----NIVSIPNKVVYLEKEKECEYLISQCK--KOSTVFVD 308	
OY	59 KYFEDTVNFVYENTIKPLIIDLYENCVCSCRAYGQTSGSKTYTLMGSPPYGOSDIPIRG 118	
DB	309 SVFDNVAIQEBEVPQTAKPLLPHYFK-GINTVFAYAGTASGSKYYTMLDK----NQNG 362	
OY	119 IFQVAAQDIFELNIYDKNTKGAFISFYELYCGKLVDLQKRK-AWALENGKEVEYVK 177	
DB	363 IVQLSLLELFITIEKKCRNIK-VLMSFLLEYNETIRDLKEKEKTKILEVEDAEAVYS 421	
OY	178 DKLILRVLTKEELLKMIIDGLVLKRIGVNSQNDESSRSHALINT---DLKDINKNT--S 231	
DB	422 NLCEIENVNVYEQAMLLINEGVYONRKMSPTRANKVSSSHALTQIYVNEILDMMNTISTY 481	
OY	232 LGKIAFIIDLASEGADTVSQNKOTQTDGANINSLALACECPAMDSDKA---HIPFR 287	
DB	482 KAKICFVDLAGSEBSAATSNGERFK-EGSYINQSLLALANCINSLSNRNISKRYVKXR 540	
OY	288 DSELTKVLARDIFVGRKSIMIANISPTISOCEQTLNLARSRYRKRNKSTCINEDDT 347	
DB	541 DSKLTHLLKMLENCULVMYTIANTPRTSPFOESNNLTLYAFRRANIK--LCATVQND 597	
OY	348 NTERISILDSKGEEMNASIENNVYIKSNHLLSNNNKNKINGKINDKIERNNILKNSFD 407	
DB	598 NKE-----SDLEIKLIKQENL-----QKEYD 618	
OY	408 KPREGFTSTFGKYSLND---IDKIKKKKKGLINTK-----STLYN 446	
DB	619 -----TLGKYNYLKEFFFIINVQLYKKQISCYKLIENTISDNKSMELKODITMY- 670	
OY	447 DNTLNKGHHNNNNNNDDNNNNNNNNNNDDSSPMNMNINHMINNNINNNTIVANNNN 506	
DB	671 DQLVKMK-----SDERYKVSLSLDQYOEKQFINLNFDTLEKRLVYNISKVDND 722	
OY	507 NNNNNNNSHNHHLPOPYVAFTDTSDFSSLDDMNCHLNNDKSIPLHKKNLLEDNFKLKRS 566	
DB	723 NNK-----SLBEM-----IPKHENKKNKNEFLVNEK 750	
OY	567 SCDNINMKKKNNLIHAHSVGSKLTFWSYDPQKNKONTFFKSNINKMEDTYPXDIYESR 626	
DB	751 VVD-----KNV-----LNGNVAVDEN-----VVDKN 772	
OY	627 NVSNANGVALLGLKANCTHHDISTDEBNHNDKINNGVINIINSNV---NSINNSMNS 682	
DB	773 NV--LNGNVAVDEVAVDENVMV--DEVAVMD-----NMVDEVNAVDEVAVDEVAVMV 823	
OY	683 INNSGNNSISYKSNVANSOSI-----SDVIQRVVENMDTSNKGNDAIFDALSCDMNY 737	
DB	824 DENVMVVENVAVDKVASVNEITVISEKHQGLSAEGSESHNNKKND--DIEDNDKOTI 880	
OY	738 PNITNNNNNNNNNNNN--NNNNIDVENYNNRDGTNSMKLYAYNSHNLFOPDNNKNTSNIQ 796	
DB	881 KDIIHNNNSGDNDDEYQANSPEES-----DYVKEKKKKIP-I 919	
OY	797 NINTNKNOGQNAVYSNMFCHYNLNDKNLYLILDNKEQCKNJIHGCDNNITIQNREDEKK 856	
DB	920 NMETKKKRTMGTKDPIHTKPIDYN---IVGLINKEDVSNKSDYNTNKNIEKNYEKK 975	
OY	857 KKTN-FVNN-----NNIV--IVNNNGKN--NSPRMKYGLGSHITSIDMKNNKMKKNEMK 907	
DB	976 GEIYPFNNLITDQNSLITYIINNHNENSHPSPMK-----1011	
OY	908 DNEKDHHSKNNNSSSSSSSNNNTIYNNINDDTFQNDYCHNDNTFTIRKNNNTNINSN 967	

Dd		1012	-----KQVAKMLLKGNTATNFIIDDDPTIK-----NNMNS	1041
Oy		968	IYQNDDIYTTINSLNDVNSNTLLHFKEKYTTPSTNEDINYKMEGCHILDDODKD	1027
Dd		1042	-----KISDKHMKMSKNILN-----NEN-----GR-----INDSKCKR	1070
Oy		1028	NDDNNVDNNKGNVNVNNDNNVNNDNNKNNVDNNDDDDVDVFHNKKDPNNNEY	1087
Dd		1071	NNNN-NNNNNNNNNNNNNNNNNSSSGGYDGINI-----LANSNYTKR	1118
Oy		1088	LSYF-----QKVADTIINNCLSLDIS-----SMYDTKEILNNILSKRYAEKONVIK	1137
Dd		1119	LHTFGGVYSLNLNDEIKTEINKKWEKMDIHLSITDSIKHARDLKXNKJKLEN----	1174
Oy		1138	YNEDIKMSLEIDKTAAOSTYEKKVLTTL	1169
Dd		1175	-FOENIKHEKHDEV-----SLYVKKKKIKKKM	1200
RESULT 26				
ID	Q815X5	PRELIMINARY;	PRT;	2612 AA.
AC	Q815X5;			
DT	01-MAR-2003	(TREMBLrel; 23,	Created)	
DT	01-MAR-2003	(TREMBLrel; 23,	Last sequence update)	
DT	01-MAR-2003	(TREMBLrel; 23,	Last annotation update)	
DE	Hypothetical protein.			
GN	PFL0350C.			
OS	Plasmodium falciparum (isolate 3D7).			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OX	NCBI_taxid=36329;			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=3D7;			
RX	MEDLINE=2255705; PubMed=12368864;			
RA	Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,			
RA	Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,			
RA	Eisen U.A., Rutherford K., Salzberg S.L., Craig A., Kyte S.,			
RA	Chan M.-S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,			
RA	Petrea M., Allen J., Selengut J., Haft D., Mather M.W., Valdivia A.B.,			
RA	Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,			
RA	McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,			
RA	Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,			
RA	Frazer C.M., Barrett B.;			
RT	"Genome sequence of the human malaria parasite Plasmodium			
RT	falciparum."			
RL	Nature 419:498-511 (2002).			
DR	EMBL; AE014845; AAN36159.1; -.			
KW	Hypothetical protein.			
SQ	SEQUENCE 2612 AA; 311421 MW; C950B6A46E3DIC5 CRC64;			
Query Match 11.8%; Score 808; DB 5; Length 2612;				
Best Local Similarity 24.2%; Pred. No. 8e-25;				
Matches 373; Conservative 220; Mismatches 434; Indels 514; Gaps 78;				
Oy		2	NSKIRVVV--RRPLSELEKKKSDIITVKNCTLYIDEPRYKDM-----TKXI	50
Dd		1174	NEKLKOEIKERRNMIONLEEKE-----FELYL-EQKYKENFENEKSGLAKKPD	1223
Oy		51	ERHEPFDVDFPD-----TVDNFTYVENTIKPLIIDLYENGCVSCFAVGOTGSKTYT	104
Dd		1224	EENEKLOINEIGEKRLKERDNFEOK-----IYEEFRNKCXY-BESIOKKYD	1274
Oy		105	MGSQGVGGSDPGIFQYAAGDIFPLNTYDKONTGFISFY-----IYCG-----	152
Dd		1275	VLEE-----DKRMKYLIIKEQEELANTYKNIITYLIDIEE	1308
Oy		153	--KLVD-----LQKRKVAALENGK--KEVVVKDLILRV-----LTXE	188
Dd		1309	KEULYVOGEKLVNKEKMLEVEKQIELEKTYKNFKEXEENDIKITIINLSOOQEDLNKE	1368
Oy		189	-ELLKMGIVLLRKI-----GVNSQNDSSRSNAIINTIDKINKTSIGKIAFIDL	240
Dd				
Dd		1369	KENIEKEQIEKRYKLDIEREBESLNNDAKOMEBSRIKDEQLDKIKAKE--ELLNTYK	1422
Oy		241	ASGERADYVSQNK-----QQTDCGANINSLALKECIRAMSDKXHIPRDELTK	293
Dd		1427	ELKREITHIMEKEKEYNKKLEESNNKKESSLNDFVELKEYSKLQEREKKLAKEKNELQK	1486
Oy		294	VLADI FG-----KSKSIMIANISPTIS--CCQOTL-----MTLRYSRY-	331
Dd		1487	V-KDQLVNYKNSLKQKEMOQNIEKEKELDEQVTYIQIDRNLSLEAKKQFILTEKHKIK	1545
Oy		332	-----KFNKSTCI NEEDDTERISIL-----DKSGEKNASSIEN	369
Dd		1546	DTEPIEOQLKLTHEOLKNKEKSLIKEN-----EINILSKLODCRNKSGKAIISLGHEN	1601
Oy		370	VVKSNNLSNNNNNNKINRG-----KINDKIER--	397
Dd		1602	RKASTN--LERTSNAR--NRGVAKYAGKNKPEDILRRNRKKSINFYNNLEINNSIQYID	1658
Oy		398	---NNILKNSPKDK-PREGFTSTFGKYSLSLIDIK-----IKKKKK--GLINYKST--	443
Dd		1659	SMINENKNGKGLKYKNDLIISDAYVS--NVWDKAASSYNNKKMKOSKLANNDTNL	1717
Oy		444	-----LYNDGTJTKK-----HNNNNNNNDNDNN-N	469
Dd		1718	VDKTDVSVIYMDSDMKVSYSGKIKENRGISNYAHITKNGANNNDNDNDNDNNNYN	1777
Oy		470	DNNNNNNDSSSWNNKMIMNMINNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	525
Dd		1778	DNNNNNNNNN--NN-----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1826
Oy		526	FTDPSFLSDMNCHLN-----NND--KSIF	550
Dd		1827	NEKKDPLNDYQYENKAYLAKDNKDYDYNQNGFPSTLPDYMENQPFNNEEKKRPY	1886
Oy		551	LHKKLELDNIKXRSSCDNINMKKKNNLHLASHVSGSKLFMSYDPQKKNDT---	606
Dd		1887	RREKELDNHHOGKINQNNNTYNNNSYVINNIISTIKKGSFNDTLNNKNSPSRLNY	1946
Oy		607	KSINIKMEDTPEKDIYESKNNVMANGVLLGL-----NKATHHS-	648
Dd		1947	KNIEHEENN-----IWTSILNVNLTLESDDLIDIKSNFFDYKONENKKGYSMWLANEQ	2001
Oy		649	----TRD--EHNDNKI-----NNGVINIIN--NSNVNSIN-----NSNNSIN-	684
Dd		2002	KAQMTADVDEYNNNTIEKESNGNASIDINININCINNDDVNDINSINNNINDNHKI	2061
Oy		685	NSNNNSIYKSNYSNQSIDVQIRYVENEDTS-----NKONDNI--FPDAI--SCDN	734
Dd		2062	NSNYGYDMGR--NYAANSNEKMYKINKKNESSLNTIDNLVANKNOQSINHSSLSMQSKHN	2120
Oy		735	NMYP-----NITNN	762
Dd		2121	KMSPK	


```

QY 988 TLAFKEKYYTPTLSTNEDIYNKMEGKHRLDDQKYDNDNNVNDNNKNNVNDVN 1047
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2469 MNITYLKSHINQSSFKK-KEKEKESNFENPENNKY-SKNNKTSMTKFLSNKSHN 2526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1048 NNDVNDVNDNDKNN-----VDNNVNDVDDDDVDPHNKKNNNNNYLS---YFQ 1092
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2527 NNNNNNAFNEKQNTNVRKIKGFMKGIOONNKIPND-----FK-FNDEYNMRKKNYE 2577
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1093 KNDVTINNCLNSLDISSMYDDTKEILN-NILSKYAEKD 1132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2578 NKHEINSNNIIEQLD-----KTLNKKSTIMNKKKKFD 2610
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 27
Q8SSM3 PRELIMINARY; PRT: 1918 AA.
ID Q8SSM3 AC Q8SSM3; DT 01-JUN-2002 (Tremblrel, 21, Created)
DT 01-JUN-2002 (Tremblrel, 21, Last sequence update)
DT 01-OCT-2002 (Tremblrel, 22, Last annotation update)
DE Erythrocyte membrane protein PFEWP3.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxId=44689;
RN RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RA "Sequence and Analysis of Chromosome 2 of Dictyostelium.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC115586; AAL92259.1; -; BFA078455C60DBE CRC64;
SQ SEQUENCE 1918 AA; 217043 MW; BFA078455C60DBE CRC64;

Query Match 11.8%; Score 807; DB 5; Length 1918;
Best Local Similarity 21.1%; Pred. No. 6.6e-25;
Matches 361; Conservative 204; Mismatches 468; Indels 674; Gaps 55;

QY 156 DLQKRMVALENGKEVYVVDLKLIR-----VLTEELLKMIIDGVLKIGVNSQND 210
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 29 EVIQTKEVQLRLQKEKEGNIKMEKFKRRKNOISDED-----INDLDD 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 211 ESSRSHILNI-DKDIINKTSLGKIAFIDLAGSEGDPTV---SONKQOTD----- 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 EDSNDDDDDNTNDDEDND-----KQKEETIKGFTSSKDVLDGDFYKVLVS 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 260 ----GANINRSL-LALKECIR--AMSDKNHIFRDELTKVLADIFVGKSKSIMANIS 312
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 GTPIGPSTNSLNGIPSTIQPFSTSSAHKPFRTSLSTFTPKNQOQKPFSLPLK 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 313 PTISCEQTLTLTKYSSRVKN-----FKNKSTCINEBDTTERIG--IIDSKGSE 361
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 STLTSTNTLPTSMTSNPEIKSIDPKDITHKRSTG-NIPNNSNNGSLNIIMKINISNE 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 362 MNASSISN-----VYKSNHL-----SNNNNKIRGKINDIEERNIL--KNKS 405
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 HNIISNNNNNTFPGPTVSKSLSLFGKNNNDNNNNNNNNNNNNNNNNNNNNNNNN 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 406 FDKREGFTSTFGKYSLLNDIKI-----KNKKGLIYKSLTYNDNTINKKHN 456
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 300 INNSKNIINNNGNLJTSNNNSNINISAGSNSTKNTSTNNNNNNNNNNNNNNNN 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 457 NNNNNNDNNNDNNNNNN-----NDS-----SNV----- 483
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 360 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 419
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 484 -----NNMIHNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 536
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 420 HNPLOAQNNVVRN--NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 477
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 537 DMN-----CHLNNDKSLFLHKQLR--DNIKLKNRSSCDNIMNKKCN----- 577
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 478 SNNNNNGNNNNNNNNINSNIGNNNVNGTNNNNNINSNINSNNNNKSSQPIORVPNTON 537
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 578 -----NLH----- 580
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 538 IKQOPLNHIQIQIQIPQSGVQQAPOQVPOQAPOQQAQOQQAQOQLLOPQPOQNRISPQ 597
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 581 -----LARHSVSGKLTWFSYDPQKNDKNTF 605
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 598 PQQPQOQNRISPQPOQTQAQAQQAQQAQQAQQAQQAQQAQQAQQAQQAQQAQQAQ 655
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 606 FKSININMEDNTPKDIYESRNVSNNG----- 633
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 656 ---RISPOQOQPTQINVCGRSIPOLDGINIDLDGLIPILQSGVPKQIRQLQOQVQWQ 712
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 634 -----NVL----- 637
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 713 QOQLQOQOQOQOQLLLOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 772
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 638 -----GLN---KNYTH-----DISTKD----- 651
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 773 QSQOYPOQNNIQSQPOQKGLTPPLKQVHTNEQDVVELLTEDSKSPISPNSNNNGRN 832
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 652 --ENHNDKIN-----GVINIINSNVNSINNSNNSINNSNNSNNSIYKSNVNS 700
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 833 IEEVNGNFTYVGNHNRGRKSGPIN--NNNINN--NNINNNNNNNNNNNNSSSSSN 889
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 701 NQSIDVQIRYVEMTSTKNDNIFPDALSCNNNTPV-----TNNNNNNNNNN 751
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 890 NIS-SNNSNNI-SNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 948
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 752 NNNNNNIDV-----NYNNDGTNSWKLY 776
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 949 NNNNNNNISNKKNNMANTPLVYKSPNNSGNNQNYINNGNNNNNNNNNNNNNNNN 1008
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 777 AYNSHILPOPNN-----KNTSNIQINNTK-----NNQGNVYSNMFCYINL-DKN-- 824
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1009 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1068
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 825 -----YLDIANKKEQDKRIHICD-----NNIIONRDEKKKNTFY-- 862
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1069 NISVNPQPSHQKILLIQAHQOQITKEHTRQOQOFLQOQIENKQSOQOQVTKHQ 1128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 863 -----NNNNIYVNNNGNNNS----- 879
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1129 HLDGLLQGHQONQPNRQIQIQPNRSGINQQLIQOYQOQOQLDQOQHLLDQOQOQOQ 1188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 880 -----PRMKYGLCGSHTSIDNNKNNEMKN 903
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1189 QPQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 1248
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 904 -----NEKNDKEMKDNHAK-----SNNNSSSSSSNNNIYNNINDDT 942
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1249 PNOQPNQPNQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO 1308
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 943 FONDYCHNDTFTIRKKNNTNINSIYONDIIYITNSLNDYNSNTLHFKEXYPTLS 1002
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1309 NNNNINHHNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1355
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1003 TNEDIYNKMEGKHRLDDQKYDNDNNVNDNNNNKNNVNDVN--DNNVNDVNDNDKNN 1061
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1356 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1402
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1062 V---DNNVNDVNDVDFPHNIRKPNNNNELYSTQKQVD-----TIIN 1100
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1403 IINNNNNNNIRNSGDGYNSTVTNNNNIIRMEYKSNNEQNPVSSKPVMI GDTSGFTLTLLN 1462
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1101 NCLNS-----LDISSYDDTKEILNLLSKYKAEKQNVIKKYINEDIKN 1146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1463 DVPKPFPEQVIALKETTFILIRIVSGAANSVYNGQTLKRLSK--AVYKGHEIDITAT 1519
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1147 SLEIDTKAQSIYKERVLLTKLLLFKKQVDTQINNETSDLRKDLVMC-----HICN 1199
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```


Db 1520 SPNDSFKLHQVKEQGVL-----QFDVKIRNGVNTPSHLVLIIFHLIDHGSIVH 1569
Qy 1200 NNPDQHFHYAVSLEKDIIILMLRQWCESEMLRLLYOFLVVEYQKXANSVLLNWS 1259
Db 1570 SIESKRISFFRSKAPYLPFIYIMQSKCFLPLSRSLYR-----NNTESNTVTNNS 1623
Qy 1260 NNGDIILLKLVODNIKSMDSMHNIH 1286
Db 1624 SN-----NNSNNNNNNNTY 1637

RESULT 28
Q81B98
ID 081B98 PRELIMINARY; PRT; 2206 AA.

AC 081B98; 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein.
GN PF07.0061.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Seeger K., Murphy L., Harris D., Berriman M., Pain A., Hall N.,
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844506; CAD50900.1; -
KW Hypothetical protein.
SQ SEQUENCE 2206 AA; 258401 MW; D4ED711AE470383C CRC64;

Query Match 11.8%; Score 804; DB 5; Length 2206;
Best Local Similarity 22.7%; Pred. No. 9,9e-25;
Matches 353; Conservative 226; Mismatches 381; Indels 598; Gaps 73;

Qy 168 ENGKKEVVVDK-LRVLTKKEELIKMIDGVLRKIGVNSQNDSSRSHALINIDKD 225
Db 467 ENVEDIIDEKEDIKEIKENVEDIIEEDIKEDIKE--NVKED-----IIEEDIKE 516
Qy 226 INKNTSLKIAFIDIASGERGADTVSONKQOTQDGANINRSLA--LKECIRAMDSD-KN 282
Db 517 -----DIKEDVKEDEIIEKENVKEDEIIEEDIKEDIKEHI--IIEEDIKE 561
Qy 283 HIPRDSLTAVLRPIFYGKSKSIMIANISPTISCEQDTLNTLRSSVAVKXKSTCIN 342
Db 562 HIIEED-----IKHHIDK-----CVEDDI--INRNISNINNSYLLN 597
Qy 343 EEDD-----TNTER-ISIDSKSEMASSIEENV-----VIKSNHLSNNNNN 384
Db 598 ENIDKIKTSKSGSHEERPRIVAKQIPANISCCPINSIEDEKYNNDLSKSTLSCSVNNN 657
Qy 385 --KINRGKINDKIERNNILKNKSPDKPREGFTSTFGYSSLANDIKIKKN--KKGII 438
Db 658 MKKKTTLKLSPESEVNNMKK-----KENETEKIKGNKIECERK-- 697
Qy 439 NYKSTLYNDNTINKGKNNNNNNNNNDNNNDNNNNNNNDSSMVMNMIMHINNINN 498
Db 698 -----NRRVSEQIDDDNNNNNNNNNNNNNNNNNNNNNNNNNNNN 735
Qy 499 INVNN 525
Db 736 NNN 795
Qy 526 -----FTDT-----SDPSSLD----- 536
Db 796 YDERKRRTFDDTIKGNLYVNDYEEVREPEESNRTKIGLQNFDFKFLPNOFILSNE 855
Qy 537 -----DMNGCLANNDSKIFL-----HKKL----- 556
Db 856 TDINENKYNIDIKSNNTKNNDDQNKVNNKNNNNNNNNNNNNNNNNNNNNNNNNNN 915
Qy 557 RDNILKLRSS-----CD-----NINMKKNNL--HLARHSGVSKLTWFSYDPQKN 603

Db 916 KDSWIKSNKKIKHATVECEDNEFLSYKKNKIKNDILOELVYKKSFRSINSINDIYNKKEKIE 975
Qy 604 TFFKSNINRKEDNTPKDI--LVESRVYNNNGVLLGLNKGTTHDSTKDENNNDKNN 661
Db 976 LNDSHLKHEVINSNNREYNOSRMHEV--VLYVMNN-----NDKDKNNERN 1025
Qy 662 GVINI--INNSVNSINN--SNMSINSNMNSIYK-----SNYSNQSISDVQIRYVN 713
Db 1026 NVCLKNKISSFVNYVMNMIMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1078
Qy 714 EMDTSKNNDNIFFDA-----ISCNNMYNITNNNNNNNNNNNNNNNNNNNNNN 757
Db 1079 ---DRNKGNIPEDMNLITNDISKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1134
Qy 758 -----IDVENYNNRDTNSM 773
Db 1135 MRAVSLSSLOKLIIEHNKILREKNNHKTQIGNODNCIFKISSESASENEENHNNNNNL 1194
Qy 774 KLY--AYNSHLFQPDNNKKT--SNIQININTKNODGNVYSMNFCYMLNDKLYLIDL 829
Db 1195 NIHSINNTNINHFPNNKKTATFVSNKTVNMN-----NITNIMNNVINSI-L 1241
Qy 830 NNKEQKDN-----IHCDNNITQN-----RNFEEKKKTTFYNNNNIYVNN-- 872
Db 1242 NNAANNMNNNAPTVMNTNSILMNNAPPTVINNTSISLMNNAPTVINNTSISLMNAP 1301
Qy 873 -NMGNNSPMMKYVLCGSHSISDMKNNEKNEMKO-----NEMKDHK 917
Db 1302 TVNINNTSNML--MNNAPTVNINNTSISLMNNAVOSOLINNGTNHMINNGTNHMMN 1358
Qy 918 S--NNNNSSSS--SSNNNIYNNI--NDDTFQ-----NDYCHNDTFT 955
Db 1359 SLINNNMMNNNIPVDIDKSKYFKNIENEPVQLNDSIQKLLHKSGPMYMNRLIS 1418
Qy 956 --IRKNNNTINSIYQN--DDIITYINSINDYMSITLLHFKEKTYFPLSTMEDLYNK 1010
Db 1419 NNLNNDKSNVNSPTFKANNVGENISTCOHNNNNYNN-----ITTSQPYNC 1465
Qy 1011 EMEKHRLDDQDKYDNDNNN--VDNNKNNVD-----NNVDN--NNVDNNVND 1058
Db 1466 -INRYNPKDYENKDKINKSNVYKVDNEKCYTIDKYGFPYNNCEVNDMKRGN 1524
Qy 1059 KANNVNNVNDDDVDVFNHKNF--NNNEYLSFYQNVDTIINNCLNSIDIS----- 1110
Db 1525 CSNT--NNKKNKDKYGVYGVYDNDNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1583
Qy 1111 --MYDTKE-----IANNILSKYKEXKNVT 1135
Db 1584 TYILIDDNDDTDISTDKRNNPLNRRNNKKNLKNLKNLKNLKNLKNLKNLKNLKN 1640
Qy 1136 KKYI-----NEDIKMSLEIDKTAQSI--YERK-----VLTKLFLLPKN 1176
Db 1641 TKYVCLGKENSRENHNNNEENINLEMIINDYHONQOVADIISTYIHSKIQYETK 1700
Qy 1177 VDTQINN-----ETSDLR-----KDLVMCHI CNNN 1201
Db 1701 KDISNNYNNVOKIOTTHGNVPHKIGDINIOENRVNVQNNNNNNNSFAENIGRLTRS 1760
Qy 1202 PDQGFHYAVSRLEKDIIILMLRQIW-----GSENLRLYQFLVVEYQ 1246
Db 1761 SSQCKDIHA--LKEIGNVQNDKSIYNNVTNMEVDEVPITCTNEN--EDK 1807
Qy 1247 NKSANSVLLNVSNNGDIILLKLVQDN-----IKSMOHNNHKK 1288
Db 1808 NDSANS--LREKKNGTLVAKKEEENNETQVRSVRLHNKKVKNETKOREIEKK 1863

RESULT 29
Q81B94
ID 081B94 PRELIMINARY; PRT; 8591 AA.
AC 081B94;
DT 01-MAR-2003 (Tremblrel. 23, Created)


```
Db      823 PKTEVIDKGVVLRGPVYASSYKISTVNGVSRHCHLIYDGSDAERML----- 878
Qy      237 FIDLASGERGADTVSONKOTODGANINRSLALKECIRAMSDKNIIPFDSELTLYLR 296
Db      879 -----RDTSTLG-----TFKIKPF-----SNPIPLIGSIFK--- 906
Qy      297 DIFVGSKIMIANISPTISCEOTLNTLRYSSRKPFKSTCINEEDPTTERISILD 356
Db      907 ---AGQCKIEV-----CSHDSAQFIQYPERKISLHNTNPTNTN----- 944
Qy      357 SKGSEMAASIEVNVVKSNNHLSNNNNKIRKINDKIERNNILKNSPKPREGFTST 416
Db      945 -----NNNNNADRRNDNDNNNTNNNNNNNNNNNNNNNNNNNNNNNNNNNN 982
Qy      417 FGKYSIANDIKIKKKKKGLINTYKSTLYNDNTIKKKNNNNNNNNDVNNNNNNNN 476
Db      983 SG-----STNN-----NGGNSANHTNNNNNNNNNNNNNNNNNNNNNNNN 1026
Qy      477 NDSSMVNNMINTN-MINNINNINIVNNNNNNNNNN--NNSSHNNLPQ--PNYAF---TD 528
Db      1027 NNNNSNNNNNNNNSSNNAPVSSHANNHSDNNNNNNNNNNNNNNNNNNNNNNNN 1086
Qy      529 TSDESLDDNACHLANNNDKSIIFLHKKNLRDNIKLRSSCDNINMKKKNNLHLAHSYGS 588
Db      1087 TNNFS--NNHSSYNNIINGIY---NYQNSRNLNNSSGFERTIILQDNGH---HEMWS 1137
Qy      589 KLTNFSYDPOKOKNTF-----FKSNINMK-EDNTPKDLIYERANVSNANGVLLGLN 640
Db      1138 NMA-----NONNFEALIVONLEFORVNTLTGRENNPS--LFSSRMO-----LR 1178
Qy      641 KNTHHIDSTKDNHNDKINNGVNIINSNVNSINNSNNNSINNSNNNSYSKYNY-- 698
Db      1179 LNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1233
Qy      699 -----NSNQISDVOIRYVENEDTSKNKNDNIFPDALSCDNNMYPINTNNNNN--- 747
Db      1234 YSVNNNGRHHNNGS---IDYLNDEESDNNNNN-----NNNNNSTSTERNNSS 1281
Qy      748 -----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 788
Db      1282 VNDSCIVENHOHYHONNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1338
Qy      789 NKNTSNIGNINTNKNKODGNVYSW-----NFCYLNLDKYYLLDLANKKODKXNIGC 842
Db      1339 NNNNS--NSNSNHHGHNHYKSHSNHNYHLNVTTPYPRNVATTRFEHAYNSL 1395
Qy      843 DNIIIONRDEPEKKKKTNFYNNNNNIVYNN--NMGNNSPRMKYGLCSHTSIDMKNE 900
Db      1396 GSCFLKXKNSWQIPQLNCTLSNHNMMNPFVNCDDNMYNFIPLYL-----D 1442
Qy      901 MKNMKNEMKMDNIIKSNNSSSSSSSNNNININDDTPFQNDYCHNDT----- 953
Db      1443 QSNSTKYENNISFINNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1502
Qy      954 -FTIRKONTNINSIYONDIIYIGINSLDYSNTLLHFKEKYTYPTLSTEDLYNKM 1012
Db      1503 KFTLERR-NINICANI--NSIFNNINRDSPLDRV-----NLASJETIQAPI 1548
Qy      1013 EGHKIRLDDODKYDDNNNNVNDN--NKNNVNDNNVNDNNVNDNNNDKNNV----- 1062
Db      1549 RNMH-SLAVNKKSEKPTNDYKINIKONNIEKRSKNGKFLINNKLIDILKKRRHS 1607
Qy      1063 -----DNNVNDNDDDVDVFNIKPFNNNEYLS-YFOKVAVDTIINNCLSLDISMYDT 1115
Db      1608 ETFLYNQWSTQDEETGEPNTINTNYAAGIYENMSLKLINYINNNINNNKNTV 1667
Qy      1116 KEILNNIILSKYAEKDNVICYINEDIKNMS-----LEIDKTAOSIYEKRVLLTJKLL 1171
Db      1668 SP-TNNII-----ECNSDLKLENNANTKULNSLCSGEFNIITENNHNNN----- 1712
Qy      1172 LFKKVVDTQINNSTDRLKDLVWCHI CNNNPDQPFH-----YAYSRLKDOIINLIMLROI 1227
```

```
Db      1713 --NNNNNMNININSH-----NH1YNNNNNNNNHFNLDHSHSYNNINDEI-----I 1755
Qy      1228 WCESENRLLYQFLVVEYONKANSVLYLVNNSNN 1261
Db      1756 IGDIKTYKLLFS-----DHNNVNONSPFOSTIKANN 1785

RESULT 31
ID      081148      PRELIMINARY;      PRT;      2763 AA.
AC      081148;
DT      01-MAR-2003 (Tremblrel. 23, Created)
DT      01-MAR-2003 (Tremblrel. 23, last sequence update)
DE      01-MAR-2003 (Tremblrel. 23, last annotation update)
PT      Hypoetical protein.
GN      pf11.0326.
OS      Plasmodium falciparum (isolate 3D7).
OC      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX      NCBI_TaxID=36329;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=3D7;
RX      MEDLINE=22255705; PubMed=12368864;
RA      Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA      Carlton J.M., Pain A., Nelson K.B., Bowman S., Paulsen I.T., James K.,
RA      Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA      Chan M.-S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,
RA      Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA      Martin D.M.A., Falcitamb A.H., Fraumholz G.M., Mungall C.,
RA      McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA      Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA      Fraser C.M., Barrrell B.;
RT      "Genome sequence of the human malaria parasite Plasmodium
RT      falciparum."
RL      Nature 419:498-511(2002).
DR      EMBL; AE014840; AAN35910.1; -.
KW      Hypoetical protein.
SQ      SEQUENCE 2763 AA; 330300 MW; 23AF0D9286193774 CRC64;

Query Match      11.7%; Score 802; DB 5; Length 2763;
Best Local Similarity 24.8%; Pred. No. 1.56-24;
Matches 316; Conservative 166; Mismatches 358; Indels 434; Gaps 59;

Qy      248 DTVSOKOTODG-ANINRSLALKECIRAMSDKNIIPFDSELTLYLRDIFVGSKSI 306
Db      11 NTSNNKKNNTNNKINNTNTSYSGVKILDS-----TNPQNNKLR----- 53
Qy      307 MIANISPTISCEOTLNTLRYSSRVKPFKSTCINEEDPTTERISILDSKSEMASS 366
Db      54 ---SVSNMMSCSRHTDLCNNY-----KFHNNN--NKEBOTNIFGKAVTSNDQYNTDR 102
Qy      367 IENVVYSNNHLSNNNN--NKINRKINDKIERNNILKN--KSPDPREGFTST----- 416
Db      103 NNNRGRDKKHARKNRKRRDKKNHGYKVDGSSNNNNNNNNNNNNNNNNNNNNNN 162
Qy      417 --FGKYSIANDIDIK-----KNKKGLI-NYKSTLYNDNTINKCH 454
Db      163 DMFKENELCMNNVVRONGESFLCNKKGIMLQDKSYNKGIFPKNDKKIKINPPNNNNN 222
Qy      455 NNNNNNNNDNDNNNNNN-----NNNNSSSMVN----- 484
Db      223 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 282
Qy      485 -----NNI-----N 488
Db      283 EDHEKNSLQKKIYDETI PVGNFVQRKKNYKQGGDIKNMMVSGPYF SYTEKYKNDNN 342
Qy      489 HMINNNINNNINNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 548
Db      343 H--NNNNINNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 400
Qy      549 IFLH-----KKNLRDNK--LKNRSSCDNINMKKKNNLHLARHSVGSKLTMFSYDPOKN 600
```



```

Db 401 YFLAKYKGGKQYKQSVSTLPYNNI CTNDEKKN------ITKKQTHEYTDTN 452
Qy 601 KDNTEFKSNINKMEDNTPKDIYESSRV-SNMGNVLLGLNKHTHDISTDENHNDKI 659
Db 453 K-----SYVDMNMFINTYKKNKASATNNVPELENNRGD--- 486
Qy 660 NGVNIINNSVNSINNSNNSINNSNNSIYKSNVNSOSI SDVQIRYVENMDTSN 719
Db 487 ---ICNVNVIDIDKNDHNNNNNNNDNSKSHSDHNSNDHND------DSDN 534
Qy 720 KNDNIIFPDASCNDNYPNITNNNNNNNNNNNN--NNNIDVNNRGGTNSMCLYA 777
Db 535 DSDSN-----DDNNNNYYYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 563
Qy 778 ---YNSHNFOPDNKNTSNIOINNTKNNODGNV-NYSNMFCHYLN----DKN----- 824
Db 584 SVEENKVIIDIMNKKKFNIN-EKSNETININHTYDVEIDIDHVDKNNIFF 642
Qy 825 ---YLIDLN-----KEOKKNI---HCCDNNIIQNRN-----DPEKK 856
Db 643 DHNSNFVDDFNMLHVNKHNHEEYKNIKLYNRFEMENKTKGKENVYIQEYEEYEDE 702
Qy 857 KKNFYNNNNIVYNNMKNNSPRKYGCG---SHSIDMKNNEMKNNEMK 912
Db 703 DLSNNYVMD--YLYNDQIQNNDSQIHDINNYDIFILSI PMNNNDNNNDNNNDNNND-- 757
Qy 913 DNHKSNNNS---SSSSSSNNNINNDNDDTFQNDY---CHNDYTFIRKNNNTIN 965
Db 758 -NNNNNNNNNDNYSSSSSKNNNYN-----YQRYIPLHHDNSHT---NNNINI 806
Qy 966 SNIVQN-DIIYIINSNDYM-----SNTLHFKKRYTYPTLSTN----- 1004
Db 807 NNIEVNPPLIYHNENHFLKQCKEKQKTDNI KKHENDKYNI KSHDNNFLOEIMKNF 866
Qy 1005 --EDI--YKEMGKHIR----- 1018
Db 867 IYEDIKTNNKFEKYSIDNVVEKTHPLDYI INNDMNYVQKEDTNNNPFHDKTDKGTIOK 926
Qy 1019 ---LDDODKYDNDNNNNVNDN---NKNVNDNNVNDNNVNDNND---NNDKKNVD 1063
Db 927 IYLLINHDHINNNKNNISHNFI SDHSKNVNDNNHVEDINQGDIDKSKSKNTNITP 986
Qy 1064 NNNVNDDDVDVDFHN--IKNFNNNEYS--YFQKAVDTI-----INNCINSL 1106
Db 987 NNNKKODYVD-DNHKCI FQKDKNTNHLDIKQNSISHTI EHGKHLNKGDIINTNKN-L 1044
Qy 1107 DISSMTD-----DTK--EILANNILSK-YKAEKDNVYKTYINEDIKNM 1146
Db 1045 HASNKYDALHRECTIYIDNISYYGODWKIDYTNNKDSQCKEKKOHTTNIQLNHLVN- 1103
Qy 1147 SLEIDKTAOSIYEKKRVLLTKLLLEKKNVDTQINNETSDKKDLVMCHICNN----- 1200
Db 1104 --EKAKTTSITINNGKVNINKNMTCNNN-----NN-----MCPYNNCSNEYD 1146
Qy 1201 ---NPDQFHFYAY 1211
Db 1147 KKNKDHQFNYK 1160

```

```

RX MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Chevreach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Felwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphrey S., Jagers K., James K.D., Johnson D., Kachornou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., Mclean J., Mooney P., Moute S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitch E.,
RA Rajadream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.,
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13."
RL Nature 419:527-531 (2002).
DR EMBL, AL929355, CAD51731.1; -.
KM Hypothetical protein.
SQ SEQUENCE 1770 AA; 207567 MW; 83195F911DFC548F CRC64;

Query Match 11.7%; Score 801; DB 5; Length 1770;
Best Local Similarity 23.0%; Pred. No. 1.1e-24;
Matches 363; Conservative 236; Mismatches 476; Indels 506; Gaps 78;

Qy 12 RPLSELEKKKKKSDITTVKN-----CTLYIDBRYKVMTKYI 50
Db 239 RPLNDNNNNISKDDIEKKKNNINSTVYNDNTNTEINTSDHNIK-DTRLEKDMEEYI 297
Qy 51 ERHEPVDKVPDVTNFTVYENTIKPLIIDYENGCVCSCFAGTGSGKTYMLGSP 110
Db 298 KKK-----NIMVSNMKNILNLYNN-----LIYDE 322
Qy 111 YGOS--DTGCIFOYAAAGDIFPLNYDKNTYGTGIFSF-YEYICGKLYLDRKKVAAL 167
Db 323 YSENIISKKGVKKEKHIEYEBCONIHMKANESTNISIDIPCCOITIDNVADATN-EQY 381
Qy 168 ENGKKEV-----VKDKILRVLRKEELLMGIDGVLRK-IGVNSQNDSSRSRAI-- 218
Db 382 DMSQDKTYNWMQKTNNDLHYHNENLFLRKIQDYFYQKYINIKFSVDTYDYIHL 441
Qy 219 --LNDLKDINKNTSGKIAFIYDLGAS-----BRGAD-----TVSQNKQOTQDANI 263
Db 442 EWFNKLKRIINNESN-----DFPGSISNWELEYETHDEIRKNNYISSNVAVFDGDKM 494
Qy 264 NRSLLAKECIRA-MDSKRN-----IPPROSE 290
Db 495 NNINILKQCLKNLKEGKDYICTNKKWMRFLQFLYNGGPCIKRISNNIYTFIPISND 554
Qy 291 L-----TKVLARDI-----VGKSKI-----MI 308
Db 555 INNNNIMVLESRYIKNLSFLNYIDHTKFIYNEPKNEHTLYKNEVYNDNDKYTHDYTL 614
Qy 309 ANISPTISCC-----EQTLLTRY-----SSRVNFKNKST--- 339
Db 615 EETNEKVCANHYHELQYVNLKBEKXITLYIEYDDKHINKEIIDEIKIKIKKSNMKG 674
Qy 340 ---CINEE-----DNTTERISILDSK-----GSENNAS---SI 367
Db 675 NTLISNDENFSSDSSNMVNIINAKNDKLNQKLFLENDKICANSHTSSNNQTEYISL 734
Qy 368 ENVAVIKSHHLN-----NNNNKINRGKI-----NDX 394
Db 735 DN--FDADYLLNPNLSRGFPNSYKLDINTDNNENVDNNGVDSNENVDNSNEN 792
Qy 395 IERN-NILKNKSFDPKRGFTSTFGKYSLANDIDIK-----KNNKGLINYSKTYND 447
Db 793 VDNNEVDSNENVDN-NEMDNDNMVNNEN-VDSNKFIFCNKNSQGRSIIKKSSTNGSR 850
Qy 448 NTINKKANNNNNNNDNDNNNDNNNNNNNDSSSVNVMNINMNNNNNNNNNNNNNNNN 507
Db 851 RNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 904

```

```

RESULT 32
Q813A3 PRELIMINARY; PRT; 1770 AA.
AC Q813A3;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN PF10225W.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36328;
RN [1]
RP SEQUENCE FROM N.A.

```


OY 508 NNNNNNNNNHLPDPYAFDTDTSPSSLDMMCHLNNNDKSLFLHKKLRLDNITLKRSS 567
 DB 905 NNNNNNNNNNNKDN-----TSD-----NNNN-----NINKEEDKKKKT 940
 OY 568 CDNIWKKKKKHLARHSVGSCLTMFSYPOKKNQNFPPKSNINMEEDTFCILYESRN 627
 DB 941 ----NKKKEN-----EKDEBK-CKKNLNGSYE-----IYELKR 970
 OY 628 VSNNGVLLG---LNKATHTDSTKQENND-NKINGVITIIINSVNSINNSNNWSI 683
 DB 971 EFEENNNIIVYDSYNNRRINNVIDLKKNEDINKINDRNIYILSPNISANEMINNFPS 1030
 OY 684 NNS-----NMSNGIYKS--NNS-NGSISDVQIYVNMEDMSSNNNNIFRDA 729
 DB 1031 YNSSNNKKNSIPCNSSNGNDIYKSCSEYNNDEKIS--SNGYLTTSOSQGTGTD---- 1083
 OY 730 ISCNMMYPIITNN 789
 DB 1084 ---GNTGSGIYEYENDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1131
 OY 790 KNTSNIQNTNNKQNGVNVSNFCHYNLNDK-NYLID--LNKEQKDKNTHGCDN 845
 DB 1132 KKRKHAKNA---QNNTTNNRVCSNCGEQVTEKINNIIDYTHLNNIQNKNNHLKNNSS 1188
 OY 846 IIONRDEKKKKKTFYNNNNNIIVYNN--NMGNNSPPMKKGLGSGHSTINMKNNMK 903
 DB 1189 TIQNGCTI-KGEOVNVKNTNNINEEDNTTLENHKK-----DQKKKHLMK 1233
 OY 904 NEMKDNEMKDNHKSNNNNSSSSNNNNIYNNINDDPTPONDY-----CHNDTFTLR 958
 DB 1234 KKLDDIDVKGQDLKSNH-----ENKQVDENMEDTNNNSMDPOQRULISYFNQK 1287
 OY 959 KNNNTINSNIYQND-----IITYTNSLNDYNSNTLLHPKEXYTPULSTNEDIYKME 1013
 DB 1288 NKKNNISNNNNKKDDDDQGVYSNITNTNSS--LH--NSCSSSSGCGNNSLYNE-- 1340
 OY 1014 GKHIRLDQDKY---DDNDNN-----VDNNKNNVNDNNVNDNNVNDNNKNNVNN 1065
 DB 1341 ----NDISKYINFNNDNDMLKQLLVPNNSNN--NNNNNNIITINSNNNNNNNNNN 1392
 OY 1066 NVNDND-DVDFH-----NIKFNNEEYLSYFOKAVDTIINNCISLSDISSMYD 1114
 DB 1393 NFKPNBESSLWHTSMTEKQPAIINSTCYIN-----VMOCL-----SVFK 1438
 OY 1115 TREILANILSKYK-----AEKQVYIKKYINEDI-----KMSLEEDK-- 1153
 DB 1439 LIYTLNHYVYKYKAVNNSSDENENNNSFINKNFTNSIPFNIFGSNNNNKKDECL 1498
 OY 1154 -----TAOSIYKRRYL-LTKILLFKNVQOIN-NETSDLRKOLV-----MC 1195
 DB 1499 LTFSEKLFOLSKHNRKGLCVNKLINLNDKYSYLFPEYEQDCHFLVDFIHNV 1558
 OY 1196 HICNNPPDOQHFYVARSLEKDIILNLM-----ROIWC-ESBNLRILYQFLVVEYQNSA 1250
 DB 1559 KYIDSVDKNOIDYLLKKEQSIIISDLFLGLEIKITSQCEYVYITQ-----PYNYSV 1614
 OY 1251 NSVLLNVSSN-NGDIILLNKK 1270
 DB 1615 NVFKKPNENNINDNLEYPFK 1635
 RESULT 33
 O8IHP9 PRELIMINARY; PRT; 2940 AA.
 AC O8IHP9;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN PF1_0480.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3D7;
 RX MADLINE=22255705; PubMed=12368864;
 RA Gardner M.J., Hall N., Fung E., White O., Berrihan M., Hyman R.W.,
 RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyse S.,
 RA Chan M.-S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,
 RA Perce M., Allen J., Selengut U., Haft D., Mather M.W., Vaidya A.B.,
 RA Martin D.M.A., Fairhead A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
 RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
 RA Fraser C.M., Barrell B.;
 RT "Genome sequence of the human malaria parasite Plasmodium
 RT falciparum."
 RL Nature 419:498-511(2002).
 DR EMBL; AE014843; AAN36060.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 2940 AA; 348040 MW; BF8748DA051BAD2 CRC64;

Query Match 11.7%; Score 800; DB 5; Length 2940;
 Best Local Similarity 23.8%; Pred. No. 1.9e-24; Indels 478; Gaps 77;
 Matches 360; Conservative 233; Mismatches 444;

OY 132 NIYDKNTGIFISFYRYCGKLYDLQK-----RKQVA--ALENGKEVAVKDL 179
 DB 1116 NNSNSNQVCLYVNHHELY-NSNYNTIOHPSSNDVPKIOKIANINMGHEHIEK 1174
 OY 180 KILRLVTKBELILKMDIGVLRKIGVNSQDESSRSHAILNIDKDKN-NTSLGKIATF 238
 DB 1175 K-----SKEE-----TKTNEQVYRS-----INQNNITILKENEI 1204
 OY 239 DLASERGAQTQVSGKQOTQDGANINRSLALKECFRANSDKN-HIFPDSLETLYLRD 297
 DB 1205 D---ENDITLQNLNINIKDMNNMD-NITNLKN--LNNINITYPYQNTILKN--NE 1253
 OY 298 I-FVGSKSIIMIANISPTISCEQTLNTR-YSSRV---KNFKKSTCIINEED--DTNT 349
 DB 1254 IOPLANNKEV-ITGHATTHSNEININIVYKNIDQINIRKNNENNDDIINKEQKDISTN 1312
 OY 350 ER-----ISLIDSKSEMNASSIENVVYKSHLLSNNNNNKINRGKINDKI 395
 DB 1313 EQNGTFYVYTSVNGEYSILDSNPTDKYKKNLNLINISSY---NTNDHIN--NYADDI 1366
 OY 396 ERNNIILNKSFDKR-----EGFTSTGK-YSLNDIDKIKKNNKKGLINYKSTLYN 446
 DB 1367 DKTN--KDNKYEKKHDKDINIYINNINSNGKNFNINSYFDLNEKEXKKXIN--IFYH 1421
 OY 447 DNTINKHNNNNNNNN-----NDNNNDNNNNNN-----NN--NND-----SS 480
 DB 1422 DNKSLKINISBELNTNTPYKICNTINTIDNNNDHKGYELNENINNSINNNDEKITTAYLN 1481
 OY 481 SMVNNMINHMT-----NNNININIVNNN--NNNNNNNNNNNNNNNNNNNNNNNN 521
 DB 1482 RVYQNGSHDILNEDVSINTSHNKNKNDNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1538
 OY 522 PNVAFTDTSDFSSLDMMCHLNNNDKSLFLHKKLRLDNITLKRSSCDNINNKKKKHL 581
 DB 1539 -----ANSYVAKNDL-QINNNNYSYIKNEKVVHQT-----NENNIIYK-----HY 1578
 OY 582 ARHSVGSCLTMFSYPOKKN-KDNTF-----FKSNI-NKMEEDTTP-KDILYESRNVSM 631
 DB 1579 SNH-VHNYITTKOTTIOEQKNIHSDKYEKCSLPEKKNIYEPNSQNSPNSQKLPNSQNLNS 1637
 OY 632 NG-----NYVLGLNKTHTHDIST-----KQENNDKNIN 660
 DB 1638 QNLPSVPLPNSQNLPSNSPMLPNCFTYVNNQTDILISLVLYVQEOKLANDQKEMSNNNIN 1697
 OY 661 NGVINIINNSVNSINNSNNM-----SINNSMNSNSIYKSYVNSNGSISDVQI 709
 DB 1698 NTKNNNAVNEVQVNNNINVAEINILPPEKKNENNLEHNSNIYHDKRNNH---DIRT 1753

Query	Match	Similarity	Score	DB	Length	Matches	Conservative	Gap
584	EKKKKTNYNNNNIYIVANNNGANNSPMKYGLCGSHISIDMKCNKKNEMKDNEMKD	913	11.7%	DB 5	1605	22.4%	1.3e-24	75
982	NNNNNSNNNNNSIDNNNSIDNNNN-----SIDN-NNNSIDNN-----NNSIN	1024	11.7%	DB 5	1605	22.4%	1.3e-24	75
914	NHKSNNNNSSSSSSNNNNIYNNIDDDTFOYDCHDNFTFIRKNNNTIN-----SNY	969	11.7%	DB 5	1605	22.4%	1.3e-24	75
1025	NNNNSIDNNNNIISSNNNNISNTMNGRKRIK-----LKSVOITNNKKNNDLDSICIKIN	1080	11.7%	DB 5	1605	22.4%	1.3e-24	75
970	QNDITTYT-----NSLNDYMSNTLHFKEKTYT-PILTSTEDYINKEME	1013	11.7%	DB 5	1605	22.4%	1.3e-24	75
1081	NNNQDITYTVDYSSGNSAAILDKOCETDYLENV-----OYNHPNISLGLTYN----	1130	11.7%	DB 5	1605	22.4%	1.3e-24	75
1014	GKHILDDODKYDDNDNNNDNNKNDVNDVND-----NNVDNNVNDKNNVNDNNV	1067	11.7%	DB 5	1605	22.4%	1.3e-24	75
1131	GYMVCESLDCGMDNNNNNNIITANGCNIKEEPENNTCYEKKNNNNNSNNIINNINI	1190	11.7%	DB 5	1605	22.4%	1.3e-24	75
1068	DNDDDDVDPHNIKFNFNNEEYLSFOKAVDTIINCL-----NSLDISMYDITKEI	1118	11.7%	DB 5	1605	22.4%	1.3e-24	75
1191	CNSSNS-----NYSNNNSNHFT--FMSNHNIEHNECVGKNNOITIDHSEEPNNVGGKGI	1245	11.7%	DB 5	1605	22.4%	1.3e-24	75
1119	LNILISLYKAEKQNVIKKYINEDIKMSLEEDIKTAOSIYEKKRVLLTTLKLFKKVD	1178	11.7%	DB 5	1605	22.4%	1.3e-24	75
1246	RRLNGRGKQKKKD--DYPE-----ETIH--ENVAKKKSKQKATL--KKKK	1291	11.7%	DB 5	1605	22.4%	1.3e-24	75
1179	TOINNETSRLKRDLYMCHICNNPPDOPHFAYASLEKDIINTLMLRQIWCESENT--RL	1236	11.7%	DB 5	1605	22.4%	1.3e-24	75
1292	CDNNNINSHMGIMITQVQENNTGNTFN-----DINKSNTNYFRSNNIQVNEININNI	1347	11.7%	DB 5	1605	22.4%	1.3e-24	75
1237	LYQFLVEYQKNSASVLLNYSNNNGDIILINKLQVON-----IKSMQH--NNI	1285	11.7%	DB 5	1605	22.4%	1.3e-24	75
1348	LYNDISTICNNNNNNNNINNNVINGENSECTIWN--HDNHNTKNSISNDKYLSTVI	1401	11.7%	DB 5	1605	22.4%	1.3e-24	75

[illegible]


```

Db      1030 -CYODNNDNDNYIG-----DDKMM-----NNNKKKNDLHYGVFAOKLISVTNIONIF 1078
Qy      866 NIVVNNMGNNSPRMKYGLCGSHTSIDNNKNNEMKNEMKNDKEMDNHKNN----- 921
Db      1079 NYV-----NHN-----YTKFISDFTTD--KSTVFSNFDLKEN--LDROIEKNNTYTN 1122
Qy      922 -----NSSSSSSNNN-----IYNN-----INDDTF-----ONDY 947
Db      1123 FHSATMFQHSQTAKLNSNFYIKIQTEIYNNPHIHKYKPSINNEPIITERQNDL 1182
Qy      948 CHN--DNTFTIRKKNNTINSNIY-----QND-----IYTSINDYNSNT 988
Db      1183 IVTTPKLILOINKEYNNTYTHKIPIDIIPTHNTQNETTCAINKSVTFEHEGCKN 1242
Qy      989 LHFKE-----KYTPILSTMEDINYKMEGKHRLDDODKDYDDNNNNVNNNNVNDN 1044
Db      1243 --HSEBLSNKNPIEENKKEINNKSIITPLROSLGTRKKQDSFSNIAPTNKN--KNY 1299
Qy      1045 VDNNNVNDNNVNDNNKNNVNDNNVNDNDDVDYFNHKKAFNNNEY--LSY----- 1090
Db      1300 DDDNNKNNKYDDDDNNNNNNNNNNNNDDNNICVGEKKHRRHAHQDSYEVQOKDKNSK 1359
Qy      1091 ---FQKXVDITINNCLNSLDSMYDPTKEILANNILSKYKAEKDY--IKYI--MEDIK 1144
Db      1360 KGSFVNHN--VSDKSDNSLEFTKLEFD--KDFK--LYEYLIEKENTDNTVTFIQEHEM 1414
Qy      1145 NMSLEIEDTKAOSIYERKYLTLTLFLFKKNDVTOINN-----ETSDLRKDLVMC 1195
Db      1415 NNALNEIECKINS--EIRAVSDHIVDYSKNNHKNKNNNNHNSRSKNSKNINKN-- 1468
Qy      1196 HICNNNDPDQFHYAAYSRLKEDILNLMRLQWCESENRLRYQFLVVE--YONKS--ANS 1252
Db      1469 DYNNN--DMNIMVVKKKKKKNNNNNDISYLEKKRNTNLFIKIPIIKIPIKQKPKEP 1526
Qy      1253 VLVNSSNNGDIIILNKELVQDNIKSMDHNNI 1285
Db      1527 IEINVSIDIENELNKKLEKN---LDSYNI 1555

RESULT 36
081184
ID      081184 PRELIMINARY; PRT; 1922 AA.
AC      081184;
DT      01-MAR-2003 (T-EMBLrel. 23, Created)
DT      01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DE      01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE      Metacaspase-like protein.
GN      PF14_0363.
OS      Plasmodium falciparum (isolate 3D7).
OC      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX      NCBI_TaxID=36329;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=3D7;
RX      MEDLINE=22255705; PubMed=12368864;
RA      Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA      Carillon J.M., Pain A., Nelson K.B., Bowman S., Paulsen I.T., James K.,
RA      Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyse S.,
RA      Chan M.-S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,
RA      Partha M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA      Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA      McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA      Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA      Fraser C.M., Barrell B.;
RT      Genome sequence of the human malaria parasite Plasmodium
RT      falciparum.
RL      Nature 419:498-511(2002).
DR      EMBL; AEO14822; AAN36976.1; --
SQ      SQUOQUE 1922 AA; 226572 MW; 5E72815A3662B31F CRC64;

```

```

Matches 365; Conservative 221; Mismatches 498; Indels 399; Gaps 72;
Qy      1 MNSKIVVVRKRLSLEKKKKDDIITYNNNTLYIDERYVVDVTKYIERHEFIVDY 60
Db      171 INGERLNR-----YIMRQRSADWL--IRN--YLYNDLP--YEINQN-----LNKN 214
Qy      61 FDD-----TYDNF--TYVENTIKPLIIDLYENGCV-----CSCFAYGQSGKTYTMLGSP 110
Db      215 INNNAGSPFNLTAKYFENYNNK--LDEYNRNIMQYNSLNKRYRSPFNITRYNNNN 273
Qy      111 YGSDTPGLFQYAGDIFFPLNIYDKONTGIFISFEYICGLYDLLQKRKVAALENG 170
Db      274 NNNN-----VSNNNVSNNNIYNNNNNNNDNNINSFOF--SKQEGLM--KIYSNNRN 322
Qy      171 KKEVVYKDLKLR-----VLTKEELKMDGVILRLRGVNSQONDESRSHALINILKOI 226
Db      323 MNDYVKTSTNTRAADNTCAQEQINIPYNTKTYLLR--SSPTEMTFKTYSNI--LKEQ 378
Qy      227 NKRTSLGKIAFIDLAGSERGADTVSQNKQOTDGANINRSLLAKECIRAMDKNHIFP 286
Db      379 PKRSSM-----PST 387
Qy      287 RDSELTK--VLRFIVGSKS-----IMANISPTISCEQTLNTR-- 326
Db      388 NSSDPFKDFGKLKRNKRVGKKKNNKINDIYINGNIFNNKIIDNNTSNAKSEGNITLHKIYN 447
Qy      327 --YSSVVRKFKKSTCIN--EEDDTYTERISLDSKGS--ENKASSIENVVIKSHLLSN 380
Db      448 FTFSSSVKLLQKKQKQLQVTEDEKSNRAVWPLNKIITYDANAITN--DKFKSFPEN 505
Qy      381 NNNKNINRGIN--DKTERNNILKNSFPKPREGFTSTGKYSLSLDDIDIKKKKK 435
Db      506 YNHQNIIPFVSTYTNMADSTKRVSPSSFD-----VERISIP 544
Qy      436 GLINYKSTLYNDNTINKKNN--NNNNNDNNNDNNNDNNNNSSGVYNNMIINHMINN 494
Db      545 RVLR-----NGNINNNKNNINNNNNINNNNNINNNNNINNNNNINNNNNINNNNN 598
Qy      495 INNNTIVNNNNNNNNNNNS-----HNHLPOENYAFPTSDSSSLDDMNC 540
Db      599 INNNNNNNNNNNNINNNINNDVNSINIKTYNIHNNNNNNLNYSVTMSVAKNQOKLSA 658
Qy      541 HANN--NDKSIPLAKK--NLRDNITKLKRSSCD-----NIMKKKNNLILARHSVSKLT 591
Db      659 QKNSLVNDGHTHYHETLINIYQFTGGQSKCNKQVAKKEKINTQI-----SYIN 712
Qy      592 MFSYDPQKKKN--TFKFS--NINMEDNTPKDILYESSRVSNNGVLLGLNKNTH 645
Db      713 NTAAYNDYTNLPVVEFYVLSKSNLNEKEDNINEKFKDNNVDKKN----- 759
Qy      646 DISTDEN-----HNDKINAGYINIIINSNNSINNSN--MNSINNSNNNSNYSYKS- 696
Db      760 ---NDENNIIRKNNVQNNNEVI--VNNQTEQIDNKKFIKPLNNITGIDPKKKNIYNTV 814
Qy      697 -----NYSNQSISDVQIRYVNMEDTS-----NKNDNIFPDASCDNN 735
Db      815 MHINDKMDLKDGYVYNTNITPH--NYINSDDKMTNNSLHINKKADDFQMTISNDT 871
Qy      736 MYFNITNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 795
Db      872 NVCYVYVYKYYINNIIFQNNNPSN--VTNYINTINNKAKPATIDENSKL--DNNINYEYV 926
Qy      796 QNINRKNNQDGNVY-----SMNFGHYLANDKYVILDLNKK--EQQKXNIHGD 843
Db      927 TKNNIYNNVNNNVVYITKESDKNLN-----NSTNHLNDLKNVYVYINPSLYIHNNN 980
Qy      844 NNIQNR--NDEFK-----KKKTN-----FYNNNN-----IVVNNNM 874
Db      981 NNIEMNKQYTNFSPNPFIDKFEYKVBYSIOPHILNOYVYNNQVAKYNPISINILENKY 1040
Qy      875 GNNNSPRMKYGLCGSHTSIDNNKNNEMKNEMKNDKNNNNSSSSSSSNNNIY 934
Db      1041 VNOFTNQENSLYNFTTHNNNNKNNNNIDILLDKAIRNOSIYVNNNGDDETKSAEHTT 1100

```

Query Match 11.7%; Score 797.5; DB 5; Length 1922;
 Best Local Similarity 24.6%; Pred. No. 1.6e-24;


```

Qy 935 N-NINDDDTFONDYCHANDNTFTTRKKNNTINSIYQNDIITYTINSANDYMSYTLHFK 993
    |||
Db 1101 NPSIN---LKKNNY-----EDKNSN-----YHSHVSIITD--KNSIYAK 1136
    |||
Qy 994 EKYVYPTLSTNEDIYNKEMEGKHRLDODQCYDND-----NNVNDNNKNNVNDNV 1046
    |||
Db 1137 NKVAEVYTKNNE---KREDLH-KDEQYTFNNNDNNNDKNNHNNHNNHNNHNNH 1191
    |||
Qy 1047 N---NNVNDNVNDKNNVNDNNVNDNDVDFNNIKNNFNNNEYSYFQKNVDITIN-- 1100
    |||
Db 1192 NNHNNHNNHNNHNNHNNHNNHNNHNNHNNHNNHNNHNNHNNHNNHNNHNNHNN 1251
    |||
Qy 1101 ---NCLNS--LDS-----SMVDPTEKL-----NNIILSKY 1127
    |||
Db 1252 NEKNPNATEKCLDINNKKKKKKIKSVHYDMDLILSKCYVDELDTCKCINMLTKNKK 1311
    |||
Qy 1128 KAEKONVYKCY---INEDIKNMSLEE---IDKTAOSIYKERRVLLTKLL-LFKKN-VDT 1179
    |||
Db 1112 LYEQ---LKKYKHKLDKDTTKKLYDECTSEKTKG-SYTMKDSYISLLCNHETLHILN 1367
    |||
Qy 1180 QINNETSDLRKDLVWCHI CANNPDQFHFYVSRLEKDI INLIMLRQIWCESENLRLLYQ 1239
    |||
Db 1368 KSNKVSDDK-----HILSYN-----HHNNKNNNNHNNHNNHNNHNNHNNHNN 1415
    |||
Qy 1240 FLVVEYQKNSANSVLNVSNNGDIL--LAKGLVDNINIKS 1279
    |||
Db 1416 INIHNHNNKDPNSGVFLPSDNTESVYNNHNRKLCNNFNVS 1458
    |||

```

RESULT 37

```

Q94GM1 PRELIMINARY: PRT; 800 AA.
ID Q94GM1
AC Q94GM1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Kinesin-like protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=530;
RN [1]
RP SEQUENCE FROM N.A.
RA Hsing Y.-I.C., Chow T.-Y., Wu H.-P., Chao Y.-T., Liu S.-M., Lee P.-F.,
RA Su C.-L., Chen C.-S., Shaw J.-F.;
RT "Oryza sativa PAC P0431G05 genomics sequence."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC069551; AAK70904.1; -.
DR Gramene; Q94GM1; -.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
DR ATP-binding; Coiled coil; Microtubules; Motor protein.
KW SEQUENCE 800 AA; 88189 MW; D9B330375AC6F256 CRC64;
SQ

```

Query Match 11.6%; Score 796; DB 10; Length 800;
 Best Local Similarity 34.7%; Pred. No. 8.2e-25;
 Matches 207; Conservative 96; Mismatches 201; Indels 92; Gaps 18;

```

Qy 3 SKIKVVRKRPLESEKKKKDSDIIVKNNCTLYIDEPYKYVDMTKYIRHSEFIVDKVD 62
    |||
Db 198 AIKVVVRKRPLEKREKVSKEEDIIIVHSSSLTVYEPKLYDLYVKEHCEFCDAVD 257
    |||
Qy 63 DTVNDTVYENTIKPLIIDLKNGCVCSFPAYGQTSKGTYYMLGQPYGQSDPTGIFQY 122
    |||
Db 258 EGVNSDEVYRETEPIIFIPOR-TRATCFAYGQTSKGTYYM---LPUR 305
    |||
Qy 123 AAGDITFLN--LYDMDNTKGIPISEYIYCGLYDLKRRKVVALLBNGKKEVVVKDLK 180
    |||

```

```

Db 306 AADDMVRLHQPYYRNQNF-LWLSYFEIYVGLFLLSDRRGLMREDOCKQVCIVGLQ 364
    |||
Qy 181 ILRYLTKEELILMID-GVLLRKIVNSQNDSESRSHALINDLKINKTISLGKAFID 239
    |||
Db 365 EFVEY-SVVOIVKEVYIRGNAAIRSTGTANESSRSHALIQAIK--KHIIAVGKISFID 421
    |||
Qy 240 LAGESEGADVDSQNGKOTODGANINRSLLALKECISAMSDKHIIPRDELTQVLRDIF 299
    |||
Db 422 LAGESEKADITTDNDROTRIEGAINSLALKECISALALDNDQHIFRSGKLTQVLRDIF 481
    |||
Qy 300 VGSKSIIMANISPTISCEQTLNTRYSSRVNFKNSTCINEEDTNTERTISILDSKG 359
    |||
Db 482 VGNSTRVWISICISPNMGSCGHTINTLRVADRVSLS-KSGNTRKEQPTGPTIPSSSDSS 540
    |||
Qy 360 SEMNASSIENVVKSNNHLSNNNNKINKRKINDK--IENNNILNKKSPKPREGPTSTF 417
    |||
Db 541 APSYPMPIETEERIAN-----QIEKKEPVEVTS-----KAENFTSNS 577
    |||
Qy 418 GKYSIANDIDIKIKKKKKGLINYKSTLYNDNTINKHNNNNNNNNNDNN--DNNNDNNNN 475
    |||
Db 578 SMEEDRPVSMITPSYNRG-----KEENGSGLANDRERVDLNSRISYN 621
    |||
Qy 476 NNDSS--SVNNNNINNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 530
    |||
Db 622 SKPQSVGSANLQEEKVTVSPRRKAYRDXDKPERQSVYAKKDSGPEVSPGYVQQAQ 681
    |||
Qy 531 DF-----SLDD--NCHLNNDNKSTFLKKNLRDNIKL 562
    |||
Db 682 QLOQQQRPTSASQNSRSSRSEKSSCDVEIDALIEEBEALIAHREKIENTMEI 737
    |||

```

RESULT 38

```

Q9LW81 PRELIMINARY: PRT; 706 AA.
ID Q9LW81
AC Q9LW81;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similarity to kinesin heavy chain.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Strydom C., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RC STRAIN=Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:131-135 (2000).
DR EMBL; AB012247; BAB02671.1; -.
DR HSSP; P17119; 3KAR.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
DR ATP-binding; Coiled coil; Microtubules; Motor protein.
KW SEQUENCE 706 AA; 79246 MW; 059658E9569282F2 CRC64;
SQ

```

Query Match 11.6%; Score 795.5; DB 10; Length 706;
 Best Local Similarity 36.8%; Pred. No. 7.6e-25;
 Matches 203; Conservative 91; Mismatches 186; Indels 71; Gaps 16;


```

Qy 3 SKIVVVRKRLSELSEKKKSDIITVKNCTLYIDBPRYKVTMTYIERHEFIVDVFD 62
Db 168 AKIVVVRKRLSELSEKKKSDIITVKNCTLYIDBPRYKVTMTYIERHEFIVDVFD 226
Qy 63 DTUNFVYVNTIETPLIILDYENCVCSPAYGQTSGKTYTLMGSPYQSDIPGIFQY 122
Db 227 EESVNDVYRETVBPVPLIFQR-IKATCFAYGQTSGKTYTLMGSPYQSDIPGIFQY 274
Qy 123 AAGDIFFLNYDKNTKGIPISEFYCYGLYDLORRKMVALENGKEVYVYDYLKIL 182
Db 275 ASRDLMLMHTTYNQGOLFVSEFYCYGLYDLORRKMVALENGKEVYVYDYLKIL 334
Qy 183 RVLTKEELIKMID-GVLLRKIGVNSQNDSSSHALINIDL-KDINKNTS-----LGKI 235
Db 335 RV-SDTAIMELIRGATSTGTGGANBESSRHAILQALIKKSVGNGSKPRRLGKL 393
Qy 236 AFIDLASSERGADTVSONKQOTGDGAINNSLALKECIRAMSDKNIHPPROSELTQV 295
Db 394 SFIDLASSERGADTVSONKQOTGDGAINNSLALKECIRAMSDKNIHPPROSELTQV 453
Qy 296 RDIIVGSKSIIIMINISPTISCCQTLNLTLYSSRVNFK-----NKST 339
Db 454 RDSFMGSKRTVMISCISSSGSCHTLNTLYKADRVKSLSKNASKKDVSSSTNMLREST 513
Qy 340 CI-----NEEDPTNERISILDSKSGEMNASSIE--NVYIKSNHLNNNNNNKIN 387
Db 514 KIPISALPTPSNFDVVN-----EMWTEENDEFPASIEODKQKMGKGLPEPVN---- 565
Qy 388 RGTINDKIERNNILKANSFPDKREGFTSTGKYSLLNDIDIKIKKKKKGLINYKSTLYND 447
Db 566 -GMAOERIPKPTI-QMKSMDPREDM-----KKSNSDNLNALLQEEEDIVNARHKOVED 618
Qy 448 NTINKKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 500
Db 619 TNNIVKEEMLLLEADQPGQOLGYSIRLNTIISOKAGILIQONRLHPOKRLREHNL 678
Qy 501 VNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 511
Db 679 VSTTGERKSQS 689

```

RESULT 39

Q8IE15 PRELIMINARY; PRT: 1192 AA.

AC O8IE15;

DT 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DE Plasmodium falciparum asparagine-rich protein.

GN Mbl13p1.63.

OS Plasmodium falciparum (isolate 3D7).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.

OX NCBI_TaxID=36329;

RN [1]

RP SEQUENCE FROM N.A.

RA Hattie B., Lemard N., Clark L., Line A., Barron A., Corton C., Berriman M., Pain A., Hall N., Akin R., Chillingworth C., Doggett J., Omond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.

RU Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL044509; CAD52272.1;

SO SEQUENCE 1192 AA; 139166 MW; 9CEB8FDDBD0A5A46 CRC64;

Query Match 11.6%; Score 795.5; DB 5; Length 1192;

Best Local Similarity 25.5%; Pred. No. 1.2e-24;

Matches 295; Conservative 163; Mismatches 367; Indels 332; Gaps 48;

```

Db 69 SLNNNDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 113
Qy 445 YNDNTINKKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 503
Db 114 GNKGINNKKYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 173
Qy 504 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 547
Db 174 FNKTNRRNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 233
Qy 548 -SIFLHKN-----LRDNLKLRSSCDIYNNK-----KN 577
Db 234 NSNTMHSRNSVEEHLRNNSIDMNSNINNYTQOTRFSSEMEENENENKRYHTGMMN 293
Qy 578 NLHLARHSVSKLTWESYDQKKNKDTFFKSNINKKEDNTPKDILYESRVSNN----- 630
Db 294 NIFPKKYDNNSSMKRTDNKKTDTSYNNKGTIN--NDNNMMYLL--RIINININEKGS 348
Qy 631 -----MGNVLLGLNKTTHDISYDENNDKINNGVNIINNSVNSINNSM- 680
Db 349 AKKGFYTNVNNKNNLKTQNNNDNMNTJEDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 408
Qy 681 -NSINNSNNNSNISTYKSNNSOSISDVQIRYNEHDTSKN--NDNIFDAISCDNNM 736
Db 409 RNLNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 442
Qy 737 YPNTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 796
Db 443 YQNRKSNMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 488
Qy 797 NINTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 836
Db 489 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 548
Qy 837 KNTHGCDNNTIQNRNDEKKKKTNPYN-----NNNIVYNNNG--NNNSPMMKGLCG 888
Db 549 MNNYGYDDNTVYNNNN--TPSTDFPSRAVGNNNNYLNNNNNNNNSAVNNNS----- 536
Qy 889 SHTSIDNNKNNEMKNNEMKNNEMKNNEMKNNEMKNNEMKNNEMKNNEMKNNEMKNNEM 948
Db 597 --SNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 647
Qy 949 HNDNFTIRKNNNTNINSNIYOND-----DIYTTINSIDNYSNTLLEHEKERTYPTLS 1002
Db 648 -----NNQNNNEDEDDDWGELGDKYIDINSIMK-KKVIYLNQLE-----A 688
Qy 1003 TNEDIYKMEGKHIR-----LDDQKYDDNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1044
Db 689 DLNLDLSKKGNDGKNNKKKKKKKDDLFVLPHNTTLVYKKKKKKKKKKKKKKNNNTSNNNN 748
Qy 1045 VDNNNYDNNYDNNYDNNYDNNYDNNYDNNYDNNYDNNYDNNYDNNYDNNYDNNYDNN 1095
Db 749 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 808
Qy 1096 -----DTIIN-----NCLNSLDISSYDPTKEILNNIILSKYAEK 1131
Db 809 TKSTVDESVKTVDEVNNKKKDLIEDKNNNNVNDVTPKGGDDNNNNNNNNNNNNNNNN 863
Qy 1132 DNVIKK-----YINEDIKNNNS-LEEDIKTAOSIYKERV-----LTKYLL 1171
Db 864 DNKEKKEDNTVAEKPAKALYFKKKKNNSPLEMYEROVYSLNKLTVENFPITTEKMCQ 923
Qy 1172 LFKKNDVT-QINNETSDLRKDLMYGICNNPPDOGHFYA-----YSRLKDI 1218
Db 924 IMESRLNTEBIOGVAVIDKALEH-----DMSWYADLCQTLKMSPNEMKKKTS 976
Qy 1219 INLIMRQIWCESNNRLLYOFLVVEYONKSANSV-----LNVSSNNNDIIL 1266
Db 977 FEIALKKTIOGEYENLPSTFESTMEKLSDENEELSFEQOKKRLIGIYVLIOE--L 1034
Qy 1267 LNKKLVDNINKNSMDHN 1283
Db 1035 FORQIVISIVISIAHD 1051

```



```

RESULT 40
081C24 PRELIMINARY; PRT: 1437 AA.
ID 081C24;
AC 081C24;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein.
GN Mal6P1.23.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN (1)
RP SEQUENCE FROM N.A.
RA Cherevach I., Davis P., Goodhead I., Stevens K., Mungall K.,
RA Berriman M., Patil A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrett B.,
RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AL844505; CAD50294.1; -.
KW Hypothetical protein.
SQ
SEQUENCE 1437 AA; 165399 MW; FOADA05528A741DA CRC64;
Query Match 11.6%; Score 795; DB 5; Length 1437;
Best Local Similarity 24.8%; Pred. No. 1,5e-24;
Matches 311; Conservative 195; Mismatches 405; Indels 342; Gaps 56;
QY 219 LNIIDLKNTSLGKIAFIDLAGSERGADTVSONKQOTDQ-ANINSLAL----- 270
DB 308 MNFDILD--NNAVKVAFCEFTDYENK-----NKNCKMCKDGENFNSSMTSPYSGNK 360
QY 271 -----KECIRAMDSKNIIPRDSLTLYARDIFPGKSKSIMIANISFTISCEQTLTTL 325
DB 361 ILEEKSNIIYHSNNYNNNNNNNNQSTHFEALNTYKT-----CELPVLL 408
QY 326 RYSSRVKFKKSTCINEEDPTERTISILDSKSEMAASSIENVIKSNHLSNNNNK 385
DB 409 T-SERVNDNNKKCKKFNLTLPISFKNTIA-----GIMNNSNNSCMMLSKYVLNNSNTS 462
QY 386 I-----NRKIND--KIERNILKNKSF-----KPREGF 413
DB 463 IFDVTNLSTLFDMNDSKNNLTFNSCNTLEHLSFDELKNSNNNDPFISNEKSDGK 522
QY 414 TSTPKYSLSLDIDIK--KAKKGLINYKSTLYNDNTINK-----HNNNNNNNDN 464
DB 523 NLKFSQFLSSNSSNGVRGANNNDLSIVMCGLDITNDN--NKKEDNPMDDNNNNNNND 580
QY 465 NDNNNDNNNNNN-----NDSSMVNNMIMINNNINNNINNNINNNNNNNNN 510
DB 581 NNNNNNNNNNNFLAKESKKEAKTEGTSTNNIOSNENNTKPGNMKSKKKNKKAALANI 640
QY 511 NNNSHNNHLPQNYAFDTSDPSL--DMNCHLNNDKSIPLAKKRLDNILK-----N 564
DB 641 RNNSINRNLKGVNF-----SEVSLNNNEKSEVOEKKKNKKKNCKEKOGEEN 695
QY 565 RSSCCNINAKKAN-----LHLARHSVSKLTMFSYDPOKQKNTFFKSN 609
DB 696 KISNKHILNKKKKNAGDEDEVTITDINLLEKSKKQKIKNNEQHNKNNIGQNN 755
QY 610 INKMDNTPKDILYESRN--VSNMNGVNL-----LGLNKVTHHDISTPKDNHNDKIN 660
DB 756 INNNNNNN-----NSKNGRNSKKGGLINNQQTYAGNNKSELIKGEKAKESNNHNN 809
QY 661 NGVINIINN-----SNVNSINNSMNSINNSMNSNSIYKSNYSNQSISDVOIR 710
DB 810 DNLFNKNNKCSYQWKKKFTDVNVIENHISDVNNAK--SSISVQSKNNENNNN----- 861
QY 711 YVNEHDTSKKNDNIFFDALISCDNNNYFVITNNNNNNNNNNNNNNNNNNNNNNNNNN 770
DB 862 --KEKDNKKKNNNN-----NN-----NNNNNNNNNNNNNNNNNNNNNNNNNN 900
QY 771 NSMKLYAVNSHLFQPDNNKNTSNIQNI-----NTNKNQDGNVNSMNFCHYLNLDK 823

```

```

DB 901 NILNNISNKENKYL--TNQKNNGLVGMSTIISTHNSNNNERGKKY--MANNYVNSH-- 955
QY 824 NYLIDLANKKQKCK-----NIHCDDNIIIONRDEFEKKKKTNNYNNNNIYIVANNMGN 878
DB 956 -----YANKENKNNINRIVISNNNN--NINNEEDSK-----NSNDILTLNLDNTI 1003
QY 879 SPRAKYGLCSHTSIDMKNNEMKNNEMKNNEMKNNIKNNN----- 921
DB 1004 QLMNHQGLDK--DNMKNNKKDDNNIATNVLTKENKKNNNNNNIGRAITTAVGTSF 1060
QY 922 -NSSSSSSNNNIYNNINDDTFQNDYCHNDFTTLRKNNNTINSNIYONDDIITYINS 980
DB 1061 INNSNIGTEMKNFVNDIIGD-----IHEKGRDN--NSIFLETDN-----YNN 1100
QY 981 LNDVMSYTLHFKEKTYPTLSTNEDIYNKEMSK-----HRLDDQDKYDDNNNNVD 1034
DB 1101 LNKLI-----MOENENFNKINRSNVKNIIDYITWENKNNNNNNNNNNNN 1145
QY 1035 NNNKNVNDVNNNNVNDVNDNDKNNVNDNNVNDDDVDYFPHIKNFNNNEYLSPQKN 1094
DB 1146 NNNNNNSNISPDNALFNNNNYNNNGHEVLYTNMMINGKIVDFDMRMNNS--YNNK- 1200
QY 1095 VDTIINNCL-----NSLDISSMYD--DTKEILNNILSKYKAEKDVVIKYNEDIKNM 1146
DB 1201 -----NDCMFLFKNLISIGLGVNLCKENKMIESLNITFKGDNIN-----NNM 1245
QY 1147 SL-----EEDKTA--QSIYERKVLTKLLLPFKKNV--DTQINNE--TSDLRKD 1191
DB 1246 SLGINSBCLTKEDFPNCAGNNNVENKSK-----EYVISPNGINDSCDYTNMLNTI 1296
QY 1192 LVMCHICNNPDDQFHF-----YAYSRLKDIINLIMLROIWSESENTLRLL-- 1237
DB 1297 FLSGNSINENFDDANETNNKINGKGENHNSYNNVGNISITDI--NSYNNSEHFLFG 1354
QY 1238 -YQFL-----VVEYQKNSANSVILLNVSNNNGDIILNKKLVQDNITKQSMDHN 1283
DB 1355 NYNFFGNNCIHKSKYKNEIKR--LTQIKLMKEEILKYKQULMSSGINEKEPEKN 1406

```

Search completed: October 2, 2003, 16:20:40
Job time : 184 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 2, 2003, 16:28:21 ; Search time 131 Seconds
(without alignments)
2537.189 Million cell updates/sec

Title: US-10-006-780-2

Perfect score: 1288

Sequence: 1 MNSKIKVVRKRPISLEKK.....KKLVQDNKSMDDNNHHKK 1288

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1288	100.0	1351	5	0814Y0
2	19	1.5	1723	5	081224
3	18	1.4	348	5	08NM71
4	18	1.4	439	5	08T832
5	18	1.4	699	5	08NM10
6	18	1.4	815	5	08T686
7	18	1.4	1518	5	08T106
8	18	1.4	1887	5	08SS58
9	18	1.4	2994	5	0952G5
10	18	1.4	3177	5	09NA13
11	18	1.4	3619	5	09U573
12	17	1.3	104	5	08NM05
13	17	1.3	172	5	08MP24
14	17	1.3	204	5	08ST26
15	17	1.3	279	5	08T1B5
16	17	1.3	284	5	08ST23

17	17	1.3	288	5	09XY02	09xy02 dugesia_jap
18	17	1.3	293	5	08MX75	08mx75 cryptospori
19	17	1.3	300	5	08NMN2	08nmn2 dictyostei
20	17	1.3	371	5	08ICP1	08icp1 plasmodium
21	17	1.3	384	5	08TMR1	08tmr1 plasmodium
22	17	1.3	391	5	025768	025768 plasmodium
23	17	1.3	419	5	025774	025774 plasmodium
24	17	1.3	426	5	08NMN6	08nmn6 dictyostei
25	17	1.3	434	5	08T834	08t834 dictyostei
26	17	1.3	440	5	08T4W5	08t4w5 plasmodium
27	17	1.3	455	5	08MXW8	08mxw8 dugesia_jap
28	17	1.3	465	5	08T1U6	08t1u6 plasmodium
29	17	1.3	496	5	08SSX7	08ssx7 dictyostei
30	17	1.3	498	5	08MP37	08mp37 dictyostei
31	17	1.3	511	5	08SS25	08ss25 dictyostei
32	17	1.3	537	5	08T1K7	08t1k7 dictyostei
33	17	1.3	560	5	061085	061085 dictyostei
34	17	1.3	587	5	08NMN3	08nmn3 dictyostei
35	17	1.3	597	5	08MYG9	08myg9 dictyostei
36	17	1.3	608	5	094474	094474 dictyostei
37	17	1.3	614	5	08T1X4	08t1x4 dictyostei
38	17	1.3	616	5	08NMJ9	08nmj9 dictyostei
39	17	1.3	628	5	08S5Y0	08s5y0 dictyostei
40	17	1.3	634	5	08NMN9	08nmn9 dictyostei
41	17	1.3	635	5	08MYH5	08myh5 dictyostei
42	17	1.3	638	5	08T2T4	08t2t4 dictyostei
43	17	1.3	647	5	08MMY6	08mmy6 dictyostei
44	17	1.3	657	5	08MMU1	08mmu1 dictyostei
45	17	1.3	662	5	08T1P5	08t1p5 dictyostei
46	17	1.3	663	5	08NMN6	08nmn6 dictyostei
47	17	1.3	682	5	09GPR8	09gpr8 dictyostei
48	17	1.3	683	5	08MXN1	08mxn1 dictyostei
49	17	1.3	723	5	08T158	08t158 dictyostei
50	17	1.3	726	5	08MXL5	08mxl5 dictyostei
51	17	1.3	728	5	09GTC7	09gvc7 leishmania
52	17	1.3	734	5	08T807	08t807 dictyostei
53	17	1.3	740	5	08T138	08t138 dictyostei
54	17	1.3	749	5	002602	002602 plasmodium
55	17	1.3	754	5	08MMT9	08mmt9 dictyostei
56	17	1.3	756	5	08MYD6	08myd6 dictyostei
57	17	1.3	759	5	09U478	09u478 dictyostei
58	17	1.3	762	5	08NMN8	08nmn8 dictyostei
59	17	1.3	791	5	08T2U7	08t2u7 dictyostei
60	17	1.3	792	5	08I535	08i535 plasmodium
61	17	1.3	806	5	08T2A3	08t2a3 dictyostei
62	17	1.3	821	5	08MNC3	08mnc3 dictyostei
63	17	1.3	823	5	08MYH7	08myh7 dictyostei
64	17	1.3	827	5	08MP44	08mp44 dictyostei
65	17	1.3	828	5	08NMN35	08nmn35 dictyostei
66	17	1.3	833	5	08SSX5	08ssx5 dictyostei
67	17	1.3	863	5	08T1K7	08t1k7 plasmodium
68	17	1.3	868	5	08MNP7	08mnp7 dictyostei
69	17	1.3	887	5	08T1U0	08t1u0 dictyostei
70	17	1.3	887	5	08T1F4	08t1f4 dictyostei
71	17	1.3	919	5	08MMV3	08mmv3 dictyostei
72	17	1.3	927	5	09BLX2	09blx2 dictyostei
73	17	1.3	930	5	08T2J1	08t2j1 dictyostei
74	17	1.3	931	5	08T2K1	08t2k1 dictyostei
75	17	1.3	942	5	096611	096611 dictyostei
76	17	1.3	946	5	08T068	08t068 plasmodium
77	17	1.3	959	5	08T1P2	08t1p2 dictyostei
78	17	1.3	980	5	08MYI3	08myi3 dictyostei
79	17	1.3	1041	5	08MMT7	08mmt7 dictyostei
80	17	1.3	1057	5	08T1J3	08t1j3 plasmodium
81	17	1.3	1060	5	08SSQ6	08ssq6 dictyostei
82	17	1.3	1125	5	08MMQ6	08mmq6 dictyostei
83	17	1.3	1130	5	08T0Z4	08t0z4 plasmodium
84	17	1.3	1139	5	023865	023865 dictyostei
85	17	1.3	1197	5	08MP87	08mp87 dictyostei
86	17	1.3	1224	5	08TAX9	08tax9 plasmodium
87	17	1.3	1231	5	09T140	09t140 dictyostei
88	17	1.3	1297	5	08T1C8	08t1c8 plasmodium
89	17	1.3	1309	5	08T2H9	08t2h9 dictyostei

90	17	1.3	1321	5	081AR1	081ar1 plasmodium	163	17	1.3	4550	5	077336	077336 plasmodium
91	17	1.3	1338	5	023927	023927 dictyosteli	164	17	1.3	4780	5	081E54	081e54 plasmodium
92	17	1.3	1363	5	081CN3	081cn3 plasmodium	165	17	1.3	5561	5	081B46	081b46 plasmodium
93	17	1.3	1376	5	08T235	08t235 dictyosteli	166	17	1.3	8591	5	081B94	081b94 plasmodium
94	17	1.3	1407	5	08T175	08t175 dictyosteli	167	17	1.3	10061	5	081J21	081j21 plasmodium
95	17	1.3	1422	5	08MMY3	08mmy3 dictyosteli	168	17	1.2	65	5	08MKQ3	08mkq3 dictyosteli
96	17	1.3	1437	5	08ICZ4	08icz4 plasmodium	169	16	1.2	68	5	08MNB2	08mnb2 dictyosteli
97	17	1.3	1457	5	044011	044011 dictyosteli	170	16	1.2	79	5	08MNB6	08mnb6 dictyosteli
98	17	1.3	1469	5	08IKP6	08ikp6 plasmodium	171	16	1.2	86	5	08MNB9	08mnb9 dictyosteli
99	17	1.3	1477	5	08ILJ1	08ilj1 plasmodium	172	16	1.2	88	5	08MNB4	08mnb4 dictyosteli
100	17	1.3	1483	5	08T273	08t273 dictyosteli	173	16	1.2	104	5	08T225	08t225 dictyosteli
101	17	1.3	1485	5	08MNB3	08mnb3 dictyosteli	174	16	1.2	130	10	08S2B4	08s2b4 oryza sativ
102	17	1.3	1502	5	08IS10	08is10 dictyosteli	175	16	1.2	135	17	08SP74	08sp74 methanosarc
103	17	1.3	1523	5	08I3B1	08i3b1 plasmodium	176	16	1.2	150	10	09FH22	09fh22 arabisidopsis
104	17	1.3	1543	5	09GV71	09gv71 dictyosteli	177	16	1.2	158	5	08MNP2	08mnp2 dictyosteli
105	17	1.3	1546	5	09SP11	09sp11 dictyosteli	178	16	1.2	160	11	035058	035058 mus musculu
106	17	1.3	1548	5	09SP10	09sp10 dictyosteli	179	16	1.2	166	5	08MND0	08mnd0 dictyosteli
107	17	1.3	1565	5	08T1M2	08t1m2 dictyosteli	180	16	1.2	185	5	08MNS7	08mns7 dictyosteli
108	17	1.3	1570	5	09U0H8	09u0h8 dictyosteli	181	16	1.2	187	5	08T279	08t279 dictyosteli
109	17	1.3	1631	5	08I3H3	08i3h3 plasmodium	182	16	1.2	188	10	064868	064868 arabisidopsis
110	17	1.3	1670	5	023901	023901 dictyosteli	183	16	1.2	198	5	08MNB8	08mnb8 dictyosteli
111	17	1.3	1672	5	08T2M6	08t2m6 dictyosteli	184	16	1.2	212	5	08T1D8	08t1d8 dictyosteli
112	17	1.3	1709	5	09SPH5	09sph5 dictyosteli	185	16	1.2	229	5	08MP31	08mp31 dictyosteli
113	17	1.3	1721	5	08SSQ0	08ssq0 dictyosteli	186	16	1.2	238	5	08T2C1	08t2c1 dictyosteli
114	17	1.3	1722	5	08I1I8	08i1i8 plasmodium	187	16	1.2	252	5	08T227	08t227 dictyosteli
115	17	1.3	1728	5	08SSU4	08ssu4 dictyosteli	188	16	1.2	259	5	096550	096550 dictyosteli
116	17	1.3	1734	5	08MNM2	08mnm2 dictyosteli	189	16	1.2	261	5	08T668	08t668 dictyosteli
117	17	1.3	1736	5	09SPH7	09sp7 dictyosteli	190	16	1.2	268	5	08T662	08t662 dictyosteli
118	17	1.3	1758	5	08ICB2	08icb2 plasmodium	191	16	1.2	275	5	08SSQ4	08ssq4 dictyosteli
119	17	1.3	1785	5	08I5Z0	08i5z0 plasmodium	192	16	1.2	282	5	08MXI8	08mx18 dictyosteli
120	17	1.3	1789	5	08T2B3	08t2b3 dictyosteli	193	16	1.2	301	5	08T1N1	08t1n1 dictyosteli
121	17	1.3	1800	5	08I7P5	08i7p5 dictyosteli	194	16	1.2	302	5	08I1P7	08i1p7 plasmodium
122	17	1.3	1812	5	08I1E3	08i1e3 plasmodium	195	16	1.2	305	5	08MNB8	08mnb8 dictyosteli
123	17	1.3	1844	5	08I1E3	08i1e3 plasmodium	196	16	1.2	308	5	025778	025778 plasmodium
124	17	1.3	1880	5	08MP27	08mp27 dictyosteli	197	16	1.2	312	5	08MNB8	08mnb8 dictyosteli
125	17	1.3	1894	5	08I5I8	08i5i8 plasmodium	198	16	1.2	312	5	08MNB4	08mnb4 dictyosteli
126	17	1.3	1894	5	08I3S3	08i3s3 plasmodium	199	16	1.2	316	5	08MNF2	08mnf2 dictyosteli
127	17	1.3	1918	5	08SSW2	08ssw2 dictyosteli	200	16	1.2	317	5	097336	097336 plasmodium
128	17	1.3	1942	5	08MYF9	08myf9 dictyosteli	201	16	1.2	322	5	08T1V7	08t1v7 dictyosteli
129	17	1.3	1969	5	015763	015763 dictyosteli	202	16	1.2	323	5	08ICG7	08icg7 plasmodium
130	17	1.3	2015	5	09U5Y1	09u5y1 dictyosteli	203	16	1.2	328	10	024463	024463 zea mays (m
131	17	1.3	2016	5	08I1L1	08i1l1 plasmodium	204	16	1.2	340	5	08T2I2	08t2i2 dictyosteli
132	17	1.3	2084	5	08I1O6	08i1o6 plasmodium	205	16	1.2	341	5	08T1C1	08t1c1 dictyosteli
133	17	1.3	2184	5	08I1O6	08i1o6 plasmodium	206	16	1.2	349	5	08MNB5	08mnb5 dictyosteli
134	17	1.3	2230	5	08MNM0	08mnm0 dictyosteli	207	16	1.2	351	5	096548	096548 dictyosteli
135	17	1.3	2235	5	08I1B09	08i1b09 plasmodium	208	16	1.2	357	5	08I281	08i281 plasmodium
136	17	1.3	2273	5	08MNB3	08mnb3 dictyosteli	209	16	1.2	369	5	08T1D9	08t1d9 dictyosteli
137	17	1.3	2344	5	08SSV9	08ssv9 dictyosteli	210	16	1.2	372	5	08I1V8	08i1v8 plasmodium
138	17	1.3	2359	5	08I297	08i297 plasmodium	211	16	1.2	373	8	035905	035905 saccharomyc
139	17	1.3	2388	5	08MYG5	08myg5 dictyosteli	212	16	1.2	373	10	09SMX7	09smx7 arabisidopsis
140	17	1.3	2569	5	08I1B68	08i1b68 plasmodium	213	16	1.2	374	5	08T1K2	08t1k2 dictyosteli
141	17	1.3	2577	5	08I1A1	08i1a1 plasmodium	214	16	1.2	383	5	08MNL9	08mnl9 candida alb
142	17	1.3	2625	5	08MNP29	08mnp29 dictyosteli	215	16	1.2	398	3	09Y764	09y764 candida alb
143	17	1.3	2722	5	08MP57	08mp57 dictyosteli	216	16	1.2	399	5	08MNB4	08mnb4 dictyosteli
144	17	1.3	2752	5	08ICT8	08ict8 plasmodium	217	16	1.2	401	5	027229	027229 dictyosteli
145	17	1.3	2763	5	08I1I48	08i1i48 plasmodium	218	16	1.2	405	5	08MNA0	08mna0 dictyosteli
146	17	1.3	2763	5	08I3X5	08i3x5 plasmodium	219	16	1.2	406	5	08MND0	08mnd0 dictyosteli
147	17	1.3	2792	5	08I552	08i552 plasmodium	220	16	1.2	407	5	08T831	08t831 dictyosteli
148	17	1.3	2874	5	08ICX0	08icx0 plasmodium	221	16	1.2	414	5	08T235	08t235 dictyosteli
149	17	1.3	2964	5	08I7W7	08i7w7 dictyosteli	222	16	1.2	417	17	08P7G9	08p7g9 methanosarc
150	17	1.3	3178	5	08I2P8	08i2p8 plasmodium	223	16	1.2	421	5	08MTX3	08mtx3 thermozon
151	17	1.3	3193	5	08I590	08i590 plasmodium	224	16	1.2	423	5	08MNP6	08mnp6 dictyosteli
152	17	1.3	3207	5	08I1H00	08i1h00 plasmodium	225	16	1.2	425	5	0967M5	0967m5 dictyosteli
153	17	1.3	3265	5	08I1D63	08i1d63 plasmodium	226	16	1.2	425	5	015755	015755 dictyosteli
154	17	1.3	3268	5	08IC30	08ic30 plasmodium	227	16	1.2	425	5	08I7K4	08i7k4 dictyosteli
155	17	1.3	3334	5	08I1I59	08i1i59 plasmodium	228	16	1.2	425	5	08T828	08t828 dictyosteli
156	17	1.3	3371	5	08I1W4	08i1w4 plasmodium	229	16	1.2	426	5	012646	012646 neocallimast
157	17	1.3	3417	5	08I1I6	08i1i6 dictyosteli	230	16	1.2	428	3	08T2C5	08t2c5 dictyosteli
158	17	1.3	3432	5	095PB9	095pb9 dictyosteli	231	16	1.2	428	5	08T8C5	08t8c5 dictyosteli
159	17	1.3	3933	5	097239	097239 plasmodium	232	16	1.2	432	5	08I1R8	08i1r8 drosophila
160	17	1.3	4138	5	08I1Y3	08i1y3 plasmodium	233	16	1.2	437	5	08SSX8	08ssx8 dictyosteli
161	17	1.3	4202	5	08I2S8	08i2s8 plasmodium	234	16	1.2	437	5	08T199	08t199 dictyosteli
162	17	1.3	4466	5	08I1S4	08i1s4 plasmodium	235	16	1.2	442	5	08SSQ9	08ssq9 dictyosteli

236	16	1.2	450	5	Q8MP78	Q8mp78 dictyosteli	309	16	1.2	700	5	Q8Te16	Q8te16 dictyosteli
237	16	1.2	457	5	Q8T283	Q8t283 dictyosteli	310	16	1.2	705	5	Q8T182	Q8t182 dictyosteli
238	16	1.2	459	5	Q8I562	Q8i562 plasmodium	311	16	1.2	714	5	Q8MN30	Q8mn30 dictyosteli
239	16	1.2	463	5	Q8MMR7	Q8mmr7 dictyosteli	312	16	1.2	714	5	Q8T1U9	Q8t1u9 plasmodium
240	16	1.2	463	5	Q8MTR4	Q8mtr4 plasmodium	313	16	1.2	715	5	Q8IKY0	Q8iky0 plasmodium
241	16	1.2	474	5	Q8SR35	Q8sr35 encaphalito	314	16	1.2	718	5	Q8XYP6	Q8xyp6 dictyosteli
242	16	1.2	476	5	Q8T1Y3	Q8t1y3 dictyosteli	315	16	1.2	720	3	Q8P3M3	Q8p3m3 neurospora
243	16	1.2	490	5	Q8T2J2	Q8t2j2 dictyosteli	316	16	1.2	721	5	Q8T2I8	Q8t2i8 dictyosteli
244	16	1.2	490	5	Q8T2R4	Q8t2r4 dictyosteli	317	16	1.2	722	5	Q8T2J2	Q8t2j2 dictyosteli
245	16	1.2	498	5	Q8T1H4	Q8t1h4 dictyosteli	318	16	1.2	726	2	Q8AJB3	Q8ajb3 rumioccoccu
246	16	1.2	508	5	Q8IDS4	Q8ids4 plasmodium	319	16	1.2	733	5	Q8T2R6	Q8t2r6 dictyosteli
247	16	1.2	513	5	Q8T6C4	Q8t6c4 dictyosteli	319	16	1.2	733	5	Q8T2R6	Q8t2r6 dictyosteli
248	16	1.2	520	5	Q8T1J4	Q8t1j4 dictyosteli	320	16	1.2	743	5	Q8NGX0	Q8ngx0 dictyosteli
249	16	1.2	520	5	Q8ISN1	Q8isn1 plasmodium	321	16	1.2	747	5	Q8MMQ3	Q8mmq3 dictyosteli
250	16	1.2	522	5	Q8I6Y6	Q8i6y6 dictyosteli	322	16	1.2	757	5	Q8GSG6	Q8gsg6 dictyosteli
251	16	1.2	527	5	Q8T1V9	Q8t1v9 dictyosteli	323	16	1.2	768	5	Q8NAX6	Q8nax6 dictyosteli
252	16	1.2	530	5	Q8T2A0	Q8t2a0 dictyosteli	324	16	1.2	772	5	Q8MNT4	Q8mnt4 dictyosteli
253	16	1.2	534	5	Q4J989	Q4j989 dictyosteli	325	16	1.2	779	5	Q8IBN6	Q8ibn6 plasmodium
254	16	1.2	535	5	Q8T2U6	Q8t2u6 dictyosteli	326	16	1.2	782	5	Q8MMV7	Q8mmv7 dictyosteli
255	16	1.2	538	5	Q8SSY9	Q8ssy9 dictyosteli	327	16	1.2	786	5	Q8IDS3	Q8ids3 plasmodium
256	16	1.2	540	5	Q76771	Q76771 dictyosteli	328	16	1.2	791	5	Q00841	Q00841 dictyosteli
257	16	1.2	540	5	Q8T1Y8	Q8t1y8 dictyosteli	329	16	1.2	798	5	Q8T200	Q8t200 dictyosteli
258	16	1.2	542	5	Q8T1R3	Q8t1r3 dictyosteli	330	16	1.2	800	5	Q4J988	Q4j988 dictyosteli
259	16	1.2	547	5	Q8ID99	Q8id99 plasmodium	331	16	1.2	802	5	Q8T1R5	Q8t1r5 plasmodium
260	16	1.2	548	5	Q8T1H3	Q8t1h3 dictyosteli	332	16	1.2	803	5	Q8T2J0	Q8t2j0 dictyosteli
261	16	1.2	548	5	Q8T1S2	Q8t1s2 dictyosteli	334	16	1.2	803	5	Q8IEC5	Q8iec5 plasmodium
262	16	1.2	559	5	Q9GPT1	Q9gpt1 dictyosteli	335	16	1.2	812	5	Q8MNB3	Q8mn39 dictyosteli
263	16	1.2	559	5	Q8T1R1	Q8t1r1 dictyosteli	336	16	1.2	812	5	Q8ISD6	Q8isd6 plasmodium
264	16	1.2	565	5	Q15754	Q15754 dictyosteli	337	16	1.2	818	5	Q8ISB3	Q8isb3 plasmodium
265	16	1.2	566	5	Q2J310	Q2j310 dictyosteli	338	16	1.2	822	5	Q8MNG7	Q8mng7 dictyosteli
266	16	1.2	566	5	Q8T285	Q8t285 dictyosteli	339	16	1.2	824	5	Q00911	Q00911 dictyosteli
267	16	1.2	566	5	Q8MND0	Q8mnd0 dictyosteli	340	16	1.2	827	5	Q8IEM2	Q8iem2 plasmodium
268	16	1.2	569	2	Q9AJB2	Q9ajb2 rumioccoccu	341	16	1.2	831	5	Q15756	Q15756 dictyosteli
269	16	1.2	571	5	Q8T2A7	Q8t2a7 dictyosteli	342	16	1.2	832	5	Q8T2K2	Q8t2k2 dictyosteli
270	16	1.2	572	5	Q8MNB9	Q8mn9 dictyosteli	343	16	1.2	837	5	Q8MYE9	Q8mye9 dictyosteli
271	16	1.2	575	5	Q7J3J4	Q7j3j4 plasmodium	344	16	1.2	851	5	Q94488	Q94488 dictyosteli
272	16	1.2	577	5	Q8MN28	Q8mn28 dictyosteli	345	16	1.2	852	5	Q8SSS7	Q8ss7 dictyosteli
273	16	1.2	578	5	Q8ICM3	Q8icm3 plasmodium	346	16	1.2	855	5	Q8T4K1	Q8t4k1 dictyosteli
274	16	1.2	579	5	Q6O950	Q6o950 dictyosteli	347	16	1.2	855	5	Q8SSV6	Q8ssv6 dictyosteli
275	16	1.2	585	5	Q8T1O8	Q8t1o8 plasmodium	348	16	1.2	860	5	Q8IK95	Q8ik95 plasmodium
276	16	1.2	586	5	Q8T2J6	Q8t2j6 dictyosteli	349	16	1.2	865	5	Q8T1P3	Q8t1p3 plasmodium
277	16	1.2	595	5	Q8T2Q4	Q8t2q4 dictyosteli	350	16	1.2	872	5	P90523	P90523 dictyosteli
278	16	1.2	605	5	Q8SSZ9	Q8ssz9 dictyosteli	351	16	1.2	876	5	Q8SSS4	Q8ss4 dictyosteli
279	16	1.2	609	5	Q9NGW8	Q9ngw8 dictyosteli	352	16	1.2	877	5	Q2J383	Q2j383 dictyosteli
280	16	1.2	616	5	Q8T1D3	Q8t1d3 dictyosteli	353	16	1.2	882	5	Q8T1I5	Q8t1i5 plasmodium
281	16	1.2	619	5	Q8MNB6	Q8mn6 dictyosteli	354	16	1.2	885	5	Q8SSZ7	Q8ssz7 dictyosteli
282	16	1.2	620	5	Q8T2C0	Q8t2c0 dictyosteli	355	16	1.2	893	5	Q8T7Y3	Q8t7y3 dictyosteli
283	16	1.2	628	5	Q9VSB3	Q9vsb3 drosophila	356	16	1.2	895	5	Q8T1B2	Q8t1b2 dictyosteli
284	16	1.2	628	5	Q8ILB8	Q8ilb8 plasmodium	357	16	1.2	898	5	Q8T1G7	Q8t1g7 dictyosteli
285	16	1.2	629	5	Q8T2A9	Q8t2a9 dictyosteli	358	16	1.2	901	5	Q8T8O6	Q8t8o6 dictyosteli
286	16	1.2	630	5	Q8IB88	Q8ib88 plasmodium	359	16	1.2	911	5	Q8MXP0	Q8mxp0 dictyosteli
287	16	1.2	637	5	Q8MND3	Q8mnd3 dictyosteli	360	16	1.2	919	5	Q8T1J2	Q8t1j2 dictyosteli
288	16	1.2	637	5	Q8IB92	Q8ib92 plasmodium	361	16	1.2	934	5	Q8SSV2	Q8ssv2 dictyosteli
289	16	1.2	637	5	Q8T1Z4	Q8t1z4 plasmodium	362	16	1.2	937	5	Q8MND6	Q8mnd6 dictyosteli
290	16	1.2	649	5	Q9VW88	Q9vw88 drosophila	363	16	1.2	942	5	Q15739	Q15739 dictyosteli
291	16	1.2	655	5	Q8I4R6	Q8i4r6 plasmodium	364	16	1.2	957	5	Q8MMW4	Q8mmw4 dictyosteli
292	16	1.2	665	5	Q8ICP6	Q8icp6 plasmodium	365	16	1.2	964	5	Q8T1D3	Q8t1d3 dictyosteli
293	16	1.2	666	5	Q8ICP6	Q8icp6 dictyosteli	366	16	1.2	965	5	Q8T856	Q8t856 dictyosteli
294	16	1.2	660	5	Q94485	Q94485 dictyosteli	367	16	1.2	969	5	Q9XZ11	Q9xz11 drosophila
295	16	1.2	660	5	Q00885	Q00885 dictyosteli	368	16	1.2	970	5	Q8T1I3	Q8t1i3 dictyosteli
296	16	1.2	665	5	Q8T2O8	Q8t2o8 dictyosteli	369	16	1.2	986	5	Q8MMQ8	Q8mmq8 dictyosteli
297	16	1.2	665	5	Q8I3I8	Q8i3i8 plasmodium	370	16	1.2	990	5	Q8I437	Q8i437 plasmodium
298	16	1.2	666	5	Q8IM66	Q8im66 plasmodium	371	16	1.2	992	5	Q8T1O4	Q8t1o4 dictyosteli
299	16	1.2	667	5	Q8T1O7	Q8t1o7 plasmodium	372	16	1.2	995	5	Q8T1J9	Q8t1j9 plasmodium
300	16	1.2	680	5	Q8T1I8	Q8t1i8 plasmodium	373	16	1.2	1004	5	Q8MP26	Q8mp26 dictyosteli
301	16	1.2	684	5	Q8I3M7	Q8i3m7 plasmodium	374	16	1.2	1017	5	Q8MMQ2	Q8mmq2 dictyosteli
302	16	1.2	685	5	Q8ID39	Q8id39 plasmodium	375	16	1.2	1022	5	Q8MMY9	Q8mmy9 dictyosteli
303	16	1.2	686	5	Q8MNJ2	Q8mnj2 dictyosteli	376	16	1.2	1022	5	Q8T1E1	Q8t1e1 plasmodium
304	16	1.2	686	5	Q96245	Q96245 plasmodium	377	16	1.2	1024	5	Q8MMW3	Q8mmw3 dictyosteli
305	16	1.2	687	5	Q8T2E7	Q8t2e7 plasmodium	378	16	1.2	1032	5	Q8I524	Q8i524 dictyosteli
306	16	1.2	687	5	Q8MN01	Q8mn01 dictyosteli	379	16	1.2	1033	5	Q8MND0	Q8mnd0 dictyosteli
307	16	1.2	693	5	Q8MXN4	Q8mxn4 dictyosteli	380	16	1.2	1036	5	Q8I355	Q8i355 plasmodium
308	16	1.2	694	5	Q8T2F2	Q8t2f2 dictyosteli	381	16	1.2	1037	5	Q8IS18	Q8is18 dictyosteli

382	16	1.2	1043	5	OBMNB7	OBmb7 dictyosteli	455	16	1.2	1364	5	OBMNB0	OBmb0 dictyosteli
383	16	1.2	1050	5	OB1106	OB1106 plasmodium	456	16	1.2	1368	5	Q9V6J0	Q9V6J0 drosophila
384	16	1.2	1064	5	Q25773	Q25773 plasmodium	457	16	1.2	1373	5	OB13A8	OB13A8 plasmodium
385	16	1.2	1075	3	Q12221	Q12221 saccharomyc	458	16	1.2	1374	5	OB1XR9	OB1XR9 plasmodium
386	16	1.2	1080	5	OB11T2	OB11T2 dictyosteli	459	16	1.2	1377	5	OB15S6	OB15S6 plasmodium
387	16	1.2	1090	5	OBMMY2	OBMMY2 dictyosteli	460	16	1.2	1380	5	OBMMN9	OBMMN9 dictyosteli
388	16	1.2	1099	5	Q9U9K8	Q9U9K8 dictyosteli	461	16	1.2	1390	5	OB16H6	OB16H6 dictyosteli
389	16	1.2	1104	5	OBMMM1	OBMMM1 dictyosteli	462	16	1.2	1390	5	O770J3	O770J3 dictyosteli
390	16	1.2	1104	5	OB1JG4	OB1JG4 plasmodium	463	16	1.2	1393	5	OB16S9	OB16S9 dictyosteli
391	16	1.2	1108	5	OB12S2	OB12S2 dictyosteli	464	16	1.2	1394	5	OB11G6	OB11G6 plasmodium
392	16	1.2	1112	5	OBMM57	OBMM57 dictyosteli	465	16	1.2	1394	5	OB1DL6	OB1DL6 plasmodium
393	16	1.2	1118	5	OB11I4	OB11I4 dictyosteli	466	16	1.2	1402	5	O966F8	O966F8 dictyosteli
394	16	1.2	1123	5	OB1811	OB1811 dictyosteli	467	16	1.2	1404	5	OBMP67	OBMP67 dictyosteli
395	16	1.2	1126	5	OB1811	OB1811 dictyosteli	468	16	1.2	1405	5	OBMMN0	OBMMN0 dictyosteli
396	16	1.2	1140	5	OB1T70	OB1T70 dictyosteli	469	16	1.2	1419	5	OB13Z3	OB13Z3 plasmodium
397	16	1.2	1145	5	OB11S6	OB11S6 plasmodium	470	16	1.2	1431	5	OBMMN1	OBMMN1 dictyosteli
398	16	1.2	1147	5	Q95PH8	Q95PH8 dictyosteli	471	16	1.2	1432	5	OB19W4	OB19W4 dictyosteli
399	16	1.2	1149	5	OB1ILP5	OB1ILP5 plasmodium	472	16	1.2	1434	5	OB1J13	OB1J13 plasmodium
400	16	1.2	1152	5	OB1216	OB1216 dictyosteli	473	16	1.2	1434	5	OB1J77	OB1J77 plasmodium
401	16	1.2	1156	5	OB1474	OB1474 plasmodium	474	16	1.2	1457	5	OB1S20	OB1S20 dictyosteli
402	16	1.2	1157	5	O00895	O00895 plasmodium	475	16	1.2	1461	5	OBST04	OBST04 dictyosteli
403	16	1.2	1165	5	OB1B76	OB1B76 plasmodium	476	16	1.2	1464	5	OB1TF6	OB1TF6 plasmodium
404	16	1.2	1166	5	OB1ILR0	OB1ILR0 plasmodium	477	16	1.2	1464	5	OB1HS3	OB1HS3 plasmodium
405	16	1.2	1172	5	OB1B93	OB1B93 plasmodium	478	16	1.2	1472	5	OB1T17	OB1T17 dictyosteli
406	16	1.2	1176	5	OBSSV4	OBSSV4 dictyosteli	479	16	1.2	1476	5	OB1676	OB1676 dictyosteli
407	16	1.2	1185	5	OB1CL1	OB1CL1 plasmodium	480	16	1.2	1476	5	OBST66	OBST66 dictyosteli
408	16	1.2	1190	5	OB1JZ6	OB1JZ6 plasmodium	481	16	1.2	1476	5	O965D3	O965D3 dictyosteli
409	16	1.2	1191	5	Q9XYJ3	Q9XYJ3 dictyosteli	482	16	1.2	1478	5	OB1K03	OB1K03 plasmodium
410	16	1.2	1192	5	OB1E15	OB1E15 plasmodium	483	16	1.2	1484	5	OB1B67	OB1B67 plasmodium
411	16	1.2	1196	5	Q23866	Q23866 dictyosteli	484	16	1.2	1486	5	Q9XZS0	Q9XZS0 dictyosteli
412	16	1.2	1197	5	OB1391	OB1391 plasmodium	485	16	1.2	1488	5	OB1DP7	OB1DP7 plasmodium
413	16	1.2	1200	5	Q97277	Q97277 plasmodium	486	16	1.2	1492	5	OBSSU1	OBSSU1 dictyosteli
414	16	1.2	1203	5	OB1863	OB1863 dictyosteli	487	16	1.2	1509	5	OBST68	OBST68 dictyosteli
415	16	1.2	1204	5	OB1J34	OB1J34 dictyosteli	488	16	1.2	1520	5	OB1687	OB1687 dictyosteli
416	16	1.2	1212	5	Q9U0L0	Q9U0L0 plasmodium	489	16	1.2	1529	5	O9GOC2	O9GOC2 dictyosteli
417	16	1.2	1213	5	Q95PH2	Q95PH2 dictyosteli	490	16	1.2	1529	5	OB1S21	OB1S21 dictyosteli
418	16	1.2	1221	5	OB1BNO	OB1BNO plasmodium	491	16	1.2	1551	5	OB1S13	OB1S13 dictyosteli
419	16	1.2	1223	5	OBST14	OBST14 dictyosteli	492	16	1.2	1552	5	OB13N4	OB13N4 plasmodium
420	16	1.2	1225	5	O15784	O15784 dictyosteli	493	16	1.2	1556	5	OBST25	OBST25 dictyosteli
421	16	1.2	1225	5	O95P12	O95P12 dictyosteli	494	16	1.2	1558	5	OB1IL5	OB1IL5 plasmodium
422	16	1.2	1236	5	OB1CJ8	OB1CJ8 plasmodium	495	16	1.2	1559	5	OB1DC3	OB1DC3 plasmodium
423	16	1.2	1245	5	OB12R3	OB12R3 plasmodium	496	16	1.2	1560	5	O9GRX5	O9GRX5 dictyosteli
424	16	1.2	1250	5	OB1611	OB1611 plasmodium	497	16	1.2	1561	5	Q9U987	Q9U987 dictyosteli
425	16	1.2	1252	5	OB1IG0	OB1IG0 plasmodium	498	16	1.2	1587	5	OB1C25	OB1C25 plasmodium
426	16	1.2	1256	5	O25770	O25770 plasmodium	499	16	1.2	1593	5	OB1T17	OB1T17 dictyosteli
427	16	1.2	1262	5	OBMM40	OBMM40 dictyosteli	500	16	1.2	1597	5	OB1JM6	OB1JM6 plasmodium
428	16	1.2	1271	5	O15749	O15749 dictyosteli	501	16	1.2	1604	5	OBMMW7	OBMMW7 dictyosteli
429	16	1.2	1287	5	O9NGS5	O9NGS5 dictyosteli	502	16	1.2	1606	5	OB1DH6	OB1DH6 plasmodium
430	16	1.2	1289	5	OB1BA2	OB1BA2 plasmodium	503	16	1.2	1612	5	OB1FNO	OB1FNO plasmodium
431	16	1.2	1290	5	O9U6L1	O9U6L1 plasmodium	504	16	1.2	1618	5	OB1S13	OB1S13 plasmodium
432	16	1.2	1297	5	OB12U8	OB12U8 dictyosteli	505	16	1.2	1619	5	O77382	O77382 plasmodium
433	16	1.2	1298	5	Q9U0T0	Q9U0T0 plasmodium	506	16	1.2	1632	5	OB1KM7	OB1KM7 plasmodium
434	16	1.2	1301	5	OBWSK5	OBWSK5 plasmodium	507	16	1.2	1641	5	OB1KZ0	OB1KZ0 plasmodium
435	16	1.2	1301	5	OB1HQ2	OB1HQ2 plasmodium	508	16	1.2	1661	5	OB1E71	OB1E71 plasmodium
436	16	1.2	1304	5	OB1K26	OB1K26 dictyosteli	509	16	1.2	1662	5	OB1569	OB1569 plasmodium
437	16	1.2	1304	5	OB1KM2	OB1KM2 plasmodium	510	16	1.2	1681	5	OB1J39	OB1J39 plasmodium
438	16	1.2	1312	5	OBMMV0	OBMMV0 dictyosteli	511	16	1.2	1697	5	OB1JN6	OB1JN6 plasmodium
439	16	1.2	1313	5	OB1IM7	OB1IM7 dictyosteli	512	16	1.2	1699	5	O95PH3	O95PH3 dictyosteli
440	16	1.2	1318	5	O95PH4	O95PH4 dictyosteli	513	16	1.2	1702	5	O9GTV7	O9GTV7 dictyosteli
441	16	1.2	1318	5	OBMMX3	OBMMX3 dictyosteli	514	16	1.2	1702	5	OB1IM14	OB1IM14 plasmodium
442	16	1.2	1320	5	OB1LJ3	OB1LJ3 plasmodium	515	16	1.2	1728	5	OB1307	OB1307 plasmodium
443	16	1.2	1324	5	OB12N1	OB12N1 dictyosteli	516	16	1.2	1754	5	OB1201	OB1201 plasmodium
444	16	1.2	1324	5	OB1F08	OB1F08 plasmodium	517	16	1.2	1770	5	OB13A3	OB13A3 plasmodium
445	16	1.2	1326	5	OB1B98	OB1B98 plasmodium	518	16	1.2	1774	5	OB1AN8	OB1AN8 plasmodium
446	16	1.2	1331	5	OB1BF6	OB1BF6 plasmodium	519	16	1.2	1792	5	OB1D94	OB1D94 plasmodium
447	16	1.2	1336	5	OB1IFC8	OB1IFC8 plasmodium	520	16	1.2	1811	5	OB1UD3	OB1UD3 plasmodium
448	16	1.2	1337	5	OB1C23	OB1C23 plasmodium	521	16	1.2	1817	5	O96253	O96253 plasmodium
449	16	1.2	1338	5	OBMMZ3	OBMMZ3 dictyosteli	522	16	1.2	1826	5	OB12Y3	OB12Y3 plasmodium
450	16	1.2	1340	5	OB1CF1	OB1CF1 plasmodium	523	16	1.2	1828	5	OB1IS4	OB1IS4 plasmodium
451	16	1.2	1342	5	OB1218	OB1218 plasmodium	524	16	1.2	1831	5	OB15F8	OB15F8 plasmodium
452	16	1.2	1348	5	OB1234	OB1234 dictyosteli	525	16	1.2	1833	5	OB1AZ5	OB1AZ5 plasmodium
453	16	1.2	1359	5	OB16G6	OB16G6 dictyosteli	526	16	1.2	1838	5	OB1INS	OB1INS dictyosteli
454	16	1.2	1364	5	OB1223	OB1223 dictyosteli	527	16	1.2	1844	5	O96201	O96201 plasmodium

528	16	1.2	1855	5	Q9TX75	Q9c75 plasmodium	601	16	1.2	3569	5	Q8IBV8	Q8iv8 plasmodium
529	16	1.2	1855	5	Q9BHN0	Q9bhn0 plasmodium	602	16	1.2	3597	5	Q8IRI5	Q8iri5 plasmodium
530	16	1.2	1858	5	Q8IAL9	Q8ial9 plasmodium	603	16	1.2	3620	5	Q968T6	Q968t6 plasmodium
531	16	1.2	1898	5	Q8ILA3	Q8ila3 plasmodium	604	16	1.2	3633	5	Q8IHL0	Q8ihl0 dictyostel
532	16	1.2	1912	5	Q9U0H1	Q9u0h1 plasmodium	605	16	1.2	3763	5	Q8RTA1	Q8rta1 dictyostel
533	16	1.2	1916	5	Q8IBCI	Q8ibci plasmodium	606	16	1.2	3787	5	Q8IIZ6	Q8iiz6 dictyostel
534	16	1.2	1922	5	Q8I2P4	Q8i2p4 plasmodium	607	16	1.2	3844	5	Q94648	Q94648 plasmodium
535	16	1.2	1923	5	Q8IBH8	Q8ibh8 plasmodium	608	16	1.2	3848	5	Q76737	Q76737 dictyostel
536	16	1.2	1929	5	Q8IHU2	Q8ihu2 dictyostel	609	16	1.2	3848	5	Q8IDR0	Q8idr0 plasmodium
537	16	1.2	1936	5	Q8ISA6	Q8isa6 plasmodium	610	16	1.2	4095	5	Q8ICN0	Q8icn0 plasmodium
538	16	1.2	1951	5	Q8ITV5	Q8itv5 plasmodium	611	16	1.2	4226	5	Q9N9H5	Q9n9h5 plasmodium
539	16	1.2	1988	5	Q8ITR8	Q8itr8 plasmodium	612	16	1.2	4226	5	Q8IHY1	Q8ihy1 plasmodium
540	16	1.2	1989	5	Q8IAU7	Q8iau7 plasmodium	613	16	1.2	4405	5	Q8ILZ2	Q8ilz2 plasmodium
541	16	1.2	1997	5	Q8ISL6	Q8isl6 plasmodium	614	16	1.2	4431	5	Q8IJI6	Q8iji6 plasmodium
542	16	1.2	2010	5	Q8ICX5	Q8icx5 plasmodium	615	16	1.2	4493	5	Q8MPA9	Q8mpa9 dictyostel
543	16	1.2	2052	5	Q9SPH6	Q9sph6 dictyostel	616	16	1.2	4494	5	Q8IS12	Q8is12 plasmodium
544	16	1.2	2062	5	Q8IBP4	Q8ibp4 plasmodium	617	16	1.2	4530	5	Q8ITK6	Q8itk6 plasmodium
545	16	1.2	2123	5	Q9U9S7	Q9u9s7 dictyostel	618	16	1.2	4638	5	Q8IK96	Q8ik96 plasmodium
546	16	1.2	2130	5	Q8IBC6	Q8ibc6 plasmodium	619	16	1.2	4662	5	Q8ILR9	Q8ilr9 plasmodium
547	16	1.2	2147	5	Q8I1H3	Q8i1h3 plasmodium	620	16	1.2	4894	5	Q8IBJ2	Q8ibj2 plasmodium
548	16	1.2	2148	5	Q8I1N6	Q8i1n6 plasmodium	621	16	1.2	4891	5	Q77372	Q77372 plasmodium
549	16	1.2	2150	5	Q238G3	Q238g3 dictyostel	622	16	1.2	5251	5	Q8I1D4	Q8iid4 plasmodium
550	16	1.2	2159	5	Q8ST00	Q8st00 dictyostel	623	16	1.2	5439	5	Q8IS86	Q8is86 plasmodium
551	16	1.2	2190	5	Q8IBL5	Q8ibl5 plasmodium	624	16	1.2	5561	5	Q8ILC9	Q8ilc9 plasmodium
552	16	1.2	2206	5	Q8IBU8	Q8ibu8 plasmodium	625	16	1.2	5910	5	Q8IAP1	Q8iap1 plasmodium
553	16	1.2	2221	5	Q8I2S9	Q8i2s9 plasmodium	626	16	1.2	6077	5	Q8IC86	Q8ic86 plasmodium
554	16	1.2	2224	5	Q8TIY5	Q8tiy5 dictyostel	627	16	1.2	6761	5	Q8IC77	Q8ic77 plasmodium
555	16	1.2	2275	5	Q8IHV8	Q8ihv8 plasmodium	628	16	1.2	9341	5	Q8I3N9	Q8i3n9 plasmodium
556	16	1.2	2283	5	Q8IC35	Q8ic35 plasmodium	629	15	1.2	125	12	Q9PWY7	Q9py7 shope fibro
557	16	1.2	2284	5	Q8ISY7	Q8isy7 plasmodium	630	15	1.2	389	5	Q8IC70	Q8ic70 plasmodium
558	16	1.2	2333	5	Q8ILH9	Q8ilh9 plasmodium	631	15	1.2	461	6	Q9SK23	Q9sk23 macaca fasc
559	16	1.2	2379	5	Q8I1W3	Q8i1w3 plasmodium	632	15	1.2	481	5	Q8SSR3	Q8ssr3 dictyostel
560	16	1.2	2380	5	Q96266	Q96266 plasmodium	633	15	1.2	481	5	Q8I3L2	Q8i3l2 plasmodium
561	16	1.2	2399	5	Q8IBB0	Q8ibb0 plasmodium	634	15	1.2	482	5	Q8ICP2	Q8icp2 plasmodium
562	16	1.2	2404	5	Q8ICD5	Q8icd5 plasmodium	635	15	1.2	515	5	Q15912	Q15912 dictyostel
563	16	1.2	2405	5	Q8I1L2	Q8i1l2 plasmodium	636	15	1.2	563	5	Q8I2B2	Q8ib2b plasmodium
564	16	1.2	2415	5	Q8I1X1	Q8i1x1 plasmodium	637	15	1.2	593	5	Q8I2U6	Q8i2u6 plasmodium
565	16	1.2	2432	5	Q8I5I9	Q8i5i9 plasmodium	638	15	1.2	608	5	Q8MMM3	Q8mmm3 dictyostel
566	16	1.2	2454	5	Q8T2G3	Q8t2g3 dictyostel	639	15	1.2	612	5	Q8IDT5	Q8idt5 plasmodium
567	16	1.2	2467	5	Q8I1D3	Q8i1d3 plasmodium	640	15	1.2	626	5	Q9W1U4	Q9w1u4 drosophila
568	16	1.2	2472	5	Q8T2M5	Q8t2m5 dictyostel	641	15	1.2	631	5	Q8I1U1	Q8i1u1 plasmodium
569	16	1.2	2472	5	Q8I1P3	Q8i1p3 plasmodium	642	15	1.2	635	5	Q8I5P8	Q8isp8 plasmodium
570	16	1.2	2515	5	Q77365	Q77365 plasmodium	643	15	1.2	670	6	Q95LT1	Q95lt1 macaca fasc
571	16	1.2	2518	5	Q8IEH2	Q8ieh2 plasmodium	644	15	1.2	670	6	Q8NNT8	Q8nnt8 macaca fasc
572	16	1.2	2535	5	Q8I4O3	Q8i4o3 plasmodium	645	15	1.2	673	4	Q96MA2	Q96ma2 homo sapien
573	16	1.2	2578	5	Q8I3P9	Q8i3p9 plasmodium	646	15	1.2	673	4	Q8N4N8	Q8n4n8 homo sapien
574	16	1.2	2612	5	Q8I5X5	Q8i5x5 plasmodium	647	15	1.2	684	10	Q940X8	Q940x8 arabidopsis
575	16	1.2	2668	5	Q8IE42	Q8ie42 plasmodium	648	15	1.2	706	10	Q9AVP4	Q9avp4 nicotiana t
576	16	1.2	2672	5	Q8IEE4	Q8iee4 plasmodium	649	15	1.2	703	10	Q9LW81	Q9lw81 arabidopsis
577	16	1.2	2700	5	Q8ILC5	Q8ilc5 plasmodium	650	15	1.2	707	10	Q9FPA5	Q9fpa5 oryza sativ
578	16	1.2	2756	10	Q9LJ60	Q9lj60 arabidopsis	651	15	1.2	725	5	Q8I1Z4	Q8i1z4 dictyostel
579	16	1.2	2770	5	Q8ITV0	Q8itv0 plasmodium	652	15	1.2	763	5	Q8I3P2	Q8ip2 plasmodium
580	16	1.2	2773	5	Q8IBP4	Q8ibp4 plasmodium	653	15	1.2	793	5	Q8SSR7	Q8ssr7 dictyostel
581	16	1.2	2792	5	Q8I4R2	Q8i4r2 plasmodium	654	15	1.2	800	10	Q94GW1	Q94gw1 oryza sativ
582	16	1.2	2814	5	Q8IC15	Q8ic15 plasmodium	655	15	1.2	816	5	Q8RTN6	Q8rtn6 dictyostel
583	16	1.2	2843	5	Q96315	Q96315 dictyostel	656	15	1.2	817	5	Q964V6	Q964v6 dictyostel
584	16	1.2	2849	5	Q8IHY4	Q8ihy4 plasmodium	657	15	1.2	825	5	Q8I2K4	Q8i2k4 plasmodium
585	16	1.2	2897	5	Q8IC71	Q8ic71 plasmodium	658	15	1.2	941	5	Q8IATU1	Q8iatu1 plasmodium
586	16	1.2	2910	5	Q8IBY8	Q8iby8 plasmodium	659	15	1.2	961	5	Q8IC12	Q8ic12 plasmodium
587	16	1.2	2932	5	Q8I4T6	Q8i4t6 plasmodium	660	15	1.2	964	5	Q8I1G7	Q8i1g7 plasmodium
588	16	1.2	2940	5	Q8IHP9	Q8ihp9 plasmodium	661	15	1.2	1025	5	Q8IC13	Q8ic13 plasmodium
589	16	1.2	2943	5	Q8IK94	Q8ik94 plasmodium	662	15	1.2	1055	5	Q8NMN5	Q8nmn5 dictyostel
590	16	1.2	3099	5	Q8MYH0	Q8myh0 dictyostel	663	15	1.2	1104	5	Q8ISE2	Q8ise2 plasmodium
591	16	1.2	3134	5	Q8I4I4	Q8i4i4 plasmodium	664	15	1.2	1139	5	Q8IUS7	Q8ius7 plasmodium
592	16	1.2	3135	5	Q8I3B5	Q8ib3b5 plasmodium	665	15	1.2	1181	5	Q97260	Q97260 plasmodium
593	16	1.2	3209	5	Q8I5D0	Q8i5d0 plasmodium	666	15	1.2	1195	5	Q8I3S1	Q8i3s1 plasmodium
594	16	1.2	3218	5	Q8ILX0	Q8ilx0 plasmodium	667	15	1.2	1313	5	Q8IIT7	Q8iit7 plasmodium
595	16	1.2	3223	5	Q8IKJ2	Q8ikj2 plasmodium	668	15	1.2	1328	5	Q8IDY9	Q8idy9 plasmodium
596	16	1.2	3248	5	Q8ICP9	Q8icp9 plasmodium	669	15	1.2	1337	5	Q9Y008	Q9y008 plasmodium
597	16	1.2	3303	5	Q8I339	Q8i339 plasmodium	670	15	1.2	1379	5	Q8I2Z2	Q8i2z2 plasmodium
598	16	1.2	3364	5	Q8IM60	Q8im60 plasmodium	671	15	1.2	1438	5	Q8IUR8	Q8iur8 plasmodium
599	16	1.2	3468	5	Q8I1O4	Q8i1o4 plasmodium	672	15	1.2	1461	5	Q8I5H0	Q8ih5h0 plasmodium
600	16	1.2	3482	5	Q8ID46	Q8id46 plasmodium	673	15	1.2	1490	5	Q8I4I6	Q8i4i6 plasmodium

674	15	1.2	1559	5	081406	081406 plasmodium	747	14	1.1	853	5	081CM6	081cm6 plasmodium
675	15	1.2	1605	5	081B13	081b13 plasmodium	748	14	1.1	867	5	081B13	081b13 plasmodium
676	15	1.2	1655	5	024754	024754 drosophila	749	14	1.1	897	10	09S842	09s842 aradidopsis
677	15	1.2	1731	5	081BQ8	081bq8 plasmodium	750	14	1.1	949	5	077386	077386 plasmodium
678	15	1.2	1813	5	081D65	081d65 plasmodium	751	14	1.1	967	10	09FHD2	09fhd2 aradidopsis
679	15	1.2	2041	5	081BK6	081bk6 plasmodium	752	14	1.1	974	5	081ED7	081ed7 plasmodium
680	15	1.2	2075	5	081HRS	081hrs plasmodium	753	14	1.1	982	5	081D06	081d06 plasmodium
681	15	1.2	2249	5	081IY8	081iy8 plasmodium	754	14	1.1	987	10	09C9A8	09c9a8 aradidopsis
682	15	1.2	2539	5	096157	096157 plasmodium	755	14	1.1	992	5	081LE1	081le1 plasmodium
683	15	1.2	2577	5	081531	081531 plasmodium	756	14	1.1	1011	5	081NL6	081nl6 plasmodium
684	15	1.2	2706	5	097292	097292 plasmodium	757	14	1.1	1031	5	081564	081564 plasmodium
685	15	1.2	2723	5	0815X3	0815x3 plasmodium	758	14	1.1	1056	10	09C9S6	09c9s6 aradidopsis
686	15	1.2	2820	5	081M32	081m32 plasmodium	759	14	1.1	1060	10	09C9C9	09c9c9 aradidopsis
687	15	1.2	3183	5	081B11	081b11 plasmodium	760	14	1.1	1078	5	081ZV5	081zv5 plasmodium
688	15	1.2	3322	5	081K10	081k10 plasmodium	761	14	1.1	1087	10	08W1Y3	08w1y3 aradidopsis
689	15	1.2	3394	5	077384	077384 plasmodium	762	14	1.1	1093	5	081SR6	081sr6 plasmodium
690	15	1.2	4044	5	081D14	081d14 plasmodium	763	14	1.1	1109	10	09SH47	09sh47 aradidopsis
691	15	1.2	4230	5	081C31	081c31 plasmodium	764	14	1.1	1162	10	09LPQ7	09lpq7 aradidopsis
692	15	1.2	4261	5	081FP4	081fp4 plasmodium	765	14	1.1	1173	5	081EM5	081em5 plasmodium
693	15	1.2	5415	5	081E74	081e74 plasmodium	766	14	1.1	1232	5	08T821	08t821 dicyosteli
694	15	1.2	5704	5	081AT1	081at1 plasmodium	767	14	1.1	1245	5	08MNT0	08mnt0 dicyosteli
695	15	1.2	6088	5	081EN1	081en1 plasmodium	768	14	1.1	1245	5	08MNT0	08mnt0 dicyosteli
696	15	1.2	6088	5	081EN1	081en1 plasmodium	769	14	1.1	1267	5	081LD1	081ld1 plasmodium
697	14	1.1	105	5	081FM7	081fm7 plasmodium	770	14	1.1	1268	10	09LUG0	09lug0 aradidopsis
698	14	1.1	150	13	09PSH8	09ps8 xenopus lae	771	14	1.1	1288	5	081D5	081d5 aradidopsis
699	14	1.1	157	10	09AVD1	09avd1 nicotiana t	772	14	1.1	1292	10	09LDN0	09ldn0 aradidopsis
700	14	1.1	157	10	09AVD2	09avd2 nicotiana t	773	14	1.1	1313	10	08L7Y8	08l7y8 aradidopsis
701	14	1.1	158	10	09FQJ8	09fqj8 cyanophora	774	14	1.1	1315	5	09BYV2	09byv2 leishmania
702	14	1.1	168	11	035068	035068 mus muscicu	775	14	1.1	1352	5	081T55	081t55 plasmodium
703	14	1.1	205	5	019307	019307 caenorhabd	776	14	1.1	1361	5	081BV5	081bv5 plasmodium
704	14	1.1	312	10	08S9D7	08s9d7 equisetum a	777	14	1.1	1444	5	081CB9	081cb9 plasmodium
705	14	1.1	347	10	093XPF	093xf8 zea mays (m	778	14	1.1	1449	5	081LP9	081lp9 plasmodium
706	14	1.1	354	5	093XPF	093xf5 zea mays (m	779	14	1.1	1465	5	08ST07	08st07 dicyosteli
707	14	1.1	375	10	048527	048527 aradidopsis	780	14	1.1	1467	5	081AN5	081an5 plasmodium
708	14	1.1	385	5	081FP5	081fp5 plasmodium	781	14	1.1	1483	5	081T07	081t07 plasmodium
709	14	1.1	393	5	081TP0	081tp0 dicyosteli	782	14	1.1	1499	13	09DDA6	09dda6 xenopus lae
710	14	1.1	407	10	093XPF	093xf9 zea mays (m	783	14	1.1	1585	5	081LE0	081le0 aradidopsis
711	14	1.1	416	5	081752	081752 plasmodium	784	14	1.1	1662	10	023274	023274 aradidopsis
712	14	1.1	421	5	0814V3	0814v3 plasmodium	785	14	1.1	1740	5	081JK9	081jk9 plasmodium
713	14	1.1	433	5	081LH7	081lh7 plasmodium	786	14	1.1	1753	5	0815W3	0815w3 plasmodium
714	14	1.1	476	5	081FN3	081fn3 plasmodium	787	14	1.1	1816	5	097275	097275 plasmodium
715	14	1.1	477	5	08MNT7	08mnt7 dicyosteli	788	14	1.1	1838	5	081D26	081d26 plasmodium
716	14	1.1	495	5	08SUN1	08sun1 encephalito	789	14	1.1	1840	5	081ED3	081ed3 plasmodium
717	14	1.1	521	5	081D10	081d10 plasmodium	790	14	1.1	1891	5	077275	077275 drosophila
718	14	1.1	548	5	081L96	081l96 plasmodium	791	14	1.1	1946	5	097291	097291 plasmodium
719	14	1.1	551	5	076494	076494 dicyosteli	792	14	1.1	1960	5	09UOK8	09uok8 plasmodium
720	14	1.1	555	5	081J43	081j43 plasmodium	793	14	1.1	1967	10	08GVH3	08gvh3 oryza sativ
721	14	1.1	558	5	0815T1	0815t1 plasmodium	794	14	1.1	1991	5	081515	081515 plasmodium
722	14	1.1	560	5	081715	081715 plasmodium	795	14	1.1	1997	5	081AP3	081ap3 plasmodium
723	14	1.1	560	5	081516	081516 plasmodium	796	14	1.1	2010	5	096171	096171 plasmodium
724	14	1.1	576	5	081D42	081d42 plasmodium	797	14	1.1	2031	5	081IR4	081ir4 plasmodium
725	14	1.1	587	5	08RT29	08rt29 dicyosteli	798	14	1.1	2108	5	0815C7	0815c7 plasmodium
726	14	1.1	601	3	09C113	09c113 candida alb	799	14	1.1	2158	10	09LUT5	09lut5 aradidopsis
727	14	1.1	627	5	081CK3	081ck3 plasmodium	800	14	1.1	2273	5	081UJ4	081uj4 plasmodium
728	14	1.1	627	5	081LX9	081lx9 plasmodium	801	14	1.1	2290	5	081UJ4	081uj4 plasmodium
729	14	1.1	642	10	08S6G6	08s6g6 oryza sativ	802	14	1.1	2371	5	081RP2	081rp2 plasmodium
730	14	1.1	652	5	077168	077168 apis mellif	803	14	1.1	2548	5	081E95	081e95 plasmodium
731	14	1.1	668	5	060964	060964 leishmania	804	14	1.1	2558	5	0815T1	0815t1 plasmodium
732	14	1.1	677	5	081C73	081c73 plasmodium	805	14	1.1	2698	5	0813F7	0813f7 plasmodium
733	14	1.1	678	13	08UUN1	08uun1 oryzias lat	806	14	1.1	3069	5	081350	081350 plasmodium
734	14	1.1	682	5	081I98	081i98 plasmodium	807	14	1.1	3270	5	081DB2	081db2 plasmodium
735	14	1.1	687	5	081KE2	081ke2 plasmodium	808	14	1.1	3347	5	081KV0	081kv0 plasmodium
736	14	1.1	693	5	096680	096680 drosophila	809	14	1.1	5922	5	081IR2	081ir2 plasmodium
737	14	1.1	699	5	015816	015816 dicyosteli	810	14	1.1	7149	5	081W09	081w09 plasmodium
738	14	1.1	730	5	08T674	08t674 dicyosteli	811	14	1.1	138	5	08MPB1	08mpb1 plasmodium
739	14	1.1	733	10	08L491	08l491 oryza sativ	812	14	1.0	169	11	035061	035061 mus muscicu
740	14	1.1	767	10	022240	022240 aradidopsis	813	13	1.0	183	3	096UG3	096ug3 kluyveromyc
741	14	1.1	773	13	P79805	P79805 morone saxa	814	13	1.0	217	5	081FN2	081fn2 plasmodium
742	14	1.1	773	13	P79805	P79805 morone saxa	815	13	1.0	277	5	09NK49	09nk49 drosophila
743	14	1.1	777	10	Q9M0X6	Q9m0x6 aradidopsis	816	13	1.0	277	5	09NK49	09nk49 drosophila
744	14	1.1	794	5	081I27	081i27 plasmodium	817	13	1.0	291	5	08T843	08t843 dicyosteli
745	14	1.1	801	10	091G54	091g54 oryza sativ	818	13	1.0	358	5	08MPB0	08mpb0 dicyosteli
746	14	1.1	810	10	08L4U3	08l4u3 oryza sativ	819	13	1.0	367	10	09M2P9	09m2p9 aradidopsis

820	13	1.0	374	10	Q94AJ2	Q94AJ2 arabidopsis	893	13	1.0	1081	5	Q24030	Q24030 drosophila
821	13	1.0	392	10	Q92RB9	Q92RB9 lycopersico	894	13	1.0	1081	5	Q8IP19	Q8IP19 drosophila
822	13	1.0	407	5	Q8IE15	Q8IE15 plasmodium	895	13	1.0	1083	5	Q45935	Q45935 caenorhabdit
823	13	1.0	411	5	Q8IDX4	Q8IDX4 plasmodium	896	13	1.0	1111	5	Q8IJA2	Q8IJA2 plasmodium
824	13	1.0	420	5	Q8SRB7	Q8SRB7 encephalito	897	13	1.0	1139	5	Q97237	Q97237 plasmodium
825	13	1.0	423	5	Q8IRK3	Q8IRK3 plasmodium	898	13	1.0	1148	5	Q8WYD8	Q8WYD8 dictyosteli
826	13	1.0	439	10	Q93XPF6	Q93XPF6 zea mays (m	899	13	1.0	1158	5	Q8MP92	Q8MP92 dictyosteli
827	13	1.0	449	5	Q8IEB0	Q8IEB0 gardia lam	900	13	1.0	1163	5	Q8MP92	Q8MP92 dictyosteli
828	13	1.0	442	5	Q8IEE4	Q8IEE4 plasmodium	901	13	1.0	1172	5	Q8IEE4	Q8IEE4 plasmodium
829	13	1.0	445	5	Q8T2E4	Q8T2E4 dictyosteli	902	13	1.0	1191	10	Q8GZU1	Q8GZU1 lycopersico
830	13	1.0	458	5	Q9BMG7	Q9BMG7 plasmodium	903	13	1.0	1212	5	Q9XZ29	Q9XZ29 drosophila
831	13	1.0	507	5	Q8MNM5	Q8MNM5 dictyosteli	904	13	1.0	1215	5	Q8I1Q6	Q8I1Q6 plasmodium
832	13	1.0	511	5	Q8MXL4	Q8MXL4 dictyosteli	905	13	1.0	1225	5	Q8I1F3	Q8I1F3 plasmodium
833	13	1.0	517	5	Q8IUP2	Q8IUP2 plasmodium	906	13	1.0	1235	5	Q8IB22	Q8IB22 plasmodium
834	13	1.0	575	5	Q8IM39	Q8IM39 plasmodium	907	13	1.0	1280	5	Q8IC07	Q8IC07 plasmodium
835	13	1.0	592	5	Q8IDT9	Q8IDT9 plasmodium	908	13	1.0	1281	5	Q15802	Q15802 plasmodium
836	13	1.0	597	5	Q8I476	Q8I476 plasmodium	909	13	1.0	1281	5	Q8IBZ5	Q8IBZ5 plasmodium
837	13	1.0	601	5	Q8I1U8	Q8I1U8 plasmodium	910	13	1.0	1299	5	Q8IUQ0	Q8IUQ0 plasmodium
838	13	1.0	604	5	Q8IS16	Q8IS16 dictyosteli	911	13	1.0	1316	5	Q8IAS8	Q8IAS8 plasmodium
839	13	1.0	608	5	Q77331	Q77331 plasmodium	912	13	1.0	1318	5	Q8IDI3	Q8IDI3 plasmodium
840	13	1.0	647	5	Q9NH62	Q9NH62 schistosoma	913	13	1.0	1335	5	Q00886	Q00886 dictyosteli
841	13	1.0	649	5	Q8TIQ9	Q8TIQ9 dictyosteli	914	13	1.0	1335	10	Q9FME7	Q9FME7 arabidopsis
842	13	1.0	674	5	Q8IB50	Q8IB50 plasmodium	915	13	1.0	1341	5	Q8IM28	Q8IM28 plasmodium
843	13	1.0	679	5	Q9W3E3	Q9W3E3 drosophila	916	13	1.0	1342	5	Q97364	Q97364 plasmodium
844	13	1.0	684	5	Q8I2M7	Q8I2M7 plasmodium	917	13	1.0	1384	5	Q8I3B3	Q8I3B3 plasmodium
845	13	1.0	707	5	Q00910	Q00910 dictyosteli	918	13	1.0	1401	5	Q8ISD4	Q8ISD4 plasmodium
846	13	1.0	717	5	Q9VM27	Q9VM27 drosophila	919	13	1.0	1404	5	Q8IKT7	Q8IKT7 plasmodium
847	13	1.0	734	5	Q8I1Z1	Q8I1Z1 plasmodium	920	13	1.0	1411	5	Q8I277	Q8I277 plasmodium
848	13	1.0	742	5	Q8ISE7	Q8ISE7 plasmodium	921	13	1.0	1495	5	Q8WML8	Q8WML8 dictyosteli
849	13	1.0	743	5	Q8ILV4	Q8ILV4 plasmodium	922	13	1.0	1542	5	Q8I2T2	Q8I2T2 plasmodium
850	13	1.0	771	5	Q9XZJ3	Q9XZJ3 dictyosteli	923	13	1.0	1550	5	Q8ID17	Q8ID17 plasmodium
851	13	1.0	786	5	Q8T1B7	Q8T1B7 dictyosteli	924	13	1.0	1554	5	Q8IUY4	Q8IUY4 plasmodium
852	13	1.0	794	10	Q94OB8	Q94OB8 arabidopsis	925	13	1.0	1575	5	Q8SSN3	Q8SSN3 dictyosteli
853	13	1.0	796	5	Q8IENS	Q8IENS plasmodium	926	13	1.0	1594	5	Q8I2N7	Q8I2N7 plasmodium
854	13	1.0	799	10	Q9IUS1	Q9IUS1 arabidopsis	927	13	1.0	1701	5	Q8I1S0	Q8I1S0 plasmodium
855	13	1.0	818	5	Q8SS10	Q8SS10 encephalito	928	13	1.0	1706	5	Q8I2W9	Q8I2W9 plasmodium
856	13	1.0	838	5	Q8ILS1	Q8ILS1 plasmodium	929	13	1.0	1737	5	Q8IEH8	Q8IEH8 plasmodium
857	13	1.0	838	5	Q8IUB9	Q8IUB9 plasmodium	930	13	1.0	1754	5	Q8TIC5	Q8TIC5 plasmodium
858	13	1.0	842	5	Q8ISJ8	Q8ISJ8 plasmodium	931	13	1.0	1790	5	Q9BK49	Q9BK49 drosophila
859	13	1.0	857	5	Q8IBF8	Q8IBF8 plasmodium	932	13	1.0	1816	5	Q8IUL6	Q8IUL6 plasmodium
860	13	1.0	863	5	Q8ICT6	Q8ICT6 plasmodium	933	13	1.0	1830	5	Q8IIR3	Q8IIR3 plasmodium
861	13	1.0	865	6	Q95LL1	Q95LL1 macaca fasc	934	13	1.0	1893	5	Q9W4J1	Q9W4J1 drosophila
862	13	1.0	899	5	Q8MMR2	Q8MMR2 dictyosteli	935	13	1.0	1903	5	Q9UON7	Q9UON7 plasmodium
863	13	1.0	904	3	Q04195	Q04195 saccharomyc	936	13	1.0	1920	5	Q46205	Q46205 drosophila
864	13	1.0	905	5	Q8ID72	Q8ID72 plasmodium	937	13	1.0	2084	5	Q8T129	Q8T129 dictyosteli
865	13	1.0	936	5	Q8I2A7	Q8I2A7 plasmodium	938	13	1.0	2121	5	Q9GYC1	Q9GYC1 leishmania
866	13	1.0	937	5	Q8MP59	Q8MP59 dictyosteli	939	13	1.0	2269	5	Q77360	Q77360 plasmodium
867	13	1.0	938	5	Q8IHY0	Q8IHY0 plasmodium	940	13	1.0	2310	5	Q8IBZ1	Q8IBZ1 plasmodium
868	13	1.0	954	5	Q9XYL1	Q9XYL1 dictyosteli	941	13	1.0	2558	5	Q8IKF6	Q8IKF6 plasmodium
869	13	1.0	960	5	Q8IEC6	Q8IEC6 plasmodium	942	13	1.0	2573	5	Q96185	Q96185 plasmodium
870	13	1.0	975	5	Q8IAY8	Q8IAY8 plasmodium	943	13	1.0	2653	5	Q8IHX2	Q8IHX2 plasmodium
871	13	1.0	1011	5	Q8IEP4	Q8IEP4 plasmodium	944	13	1.0	2657	5	Q8ILG1	Q8ILG1 plasmodium
872	13	1.0	1016	5	Q8IEB3	Q8IEB3 plasmodium	945	13	1.0	2691	5	Q8ILS2	Q8ILS2 plasmodium
873	13	1.0	1029	5	Q8IKV7	Q8IKV7 plasmodium	946	13	1.0	2705	5	Q8ISL0	Q8ISL0 plasmodium
874	13	1.0	1032	10	Q8IFJ9	Q8IFJ9 arabidopsis	947	13	1.0	3001	5	Q8IL08	Q8IL08 plasmodium
875	13	1.0	1033	5	Q8IEB8	Q8IEB8 plasmodium	948	13	1.0	3063	5	Q8I284	Q8I284 plasmodium
876	13	1.0	1035	10	Q8G571	Q8G571 arabidopsis	949	13	1.0	3177	5	Q8IEH5	Q8IEH5 plasmodium
877	13	1.0	1044	5	Q9G738	Q9G738 plasmodium	950	13	1.0	3429	5	Q8IBP1	Q8IBP1 plasmodium
878	13	1.0	1044	5	Q8IS70	Q8IS70 plasmodium	951	13	1.0	3535	5	Q8IC29	Q8IC29 plasmodium
879	13	1.0	1048	5	Q9VWK4	Q9VWK4 drosophila	952	13	1.0	3574	5	Q8IDA8	Q8IDA8 plasmodium
880	13	1.0	1048	5	Q9NGN7	Q9NGN7 drosophila	953	13	1.0	3610	5	Q968T7	Q968T7 plasmodium
881	13	1.0	1051	10	Q94LM7	Q94LM7 arabidopsis	954	13	1.0	3628	5	Q968Y0	Q968Y0 plasmodium
882	13	1.0	1063	5	Q8ICV5	Q8ICV5 plasmodium	955	13	1.0	3704	5	Q8IKY8	Q8IKY8 plasmodium
883	13	1.0	1065	5	Q24029	Q24029 drosophila	956	13	1.0	5767	5	Q8IS25	Q8IS25 plasmodium
884	13	1.0	1065	5	Q8IP21	Q8IP21 drosophila	957	13	1.0	5890	5	Q8IK84	Q8IK84 plasmodium
885	13	1.0	1072	5	Q9VWK6	Q9VWK6 drosophila	958	13	1.0	5987	5	Q8IDN3	Q8IDN3 plasmodium
886	13	1.0	1072	5	Q24031	Q24031 drosophila	959	13	1.0	6118	5	Q8I396	Q8I396 plasmodium
887	13	1.0	1072	5	Q24027	Q24027 drosophila	960	13	1.0	6473	5	Q8IKH9	Q8IKH9 plasmodium
888	13	1.0	1073	5	Q8IJW5	Q8IJW5 plasmodium	961	12	0.9	74	5	Q9GVR2	Q9GVR2 plasmodium
889	13	1.0	1074	5	Q24028	Q24028 drosophila	962	12	0.9	106	10	Q8IT11	Q8IT11 lycopersico
890	13	1.0	1074	5	Q8IP20	Q8IP20 drosophila	963	12	0.9	108	5	Q8T211	Q8T211 dictyosteli
891	13	1.0	1075	10	Q9SNE3	Q9SNE3 arabidopsis	964	12	0.9	125	5	Q8T2H7	Q8T2H7 dictyosteli
892	13	1.0	1076	5	Q8IL74	Q8IL74 plasmodium	965	12	0.9	155	11	Q99PR6	Q99PR6 mus musculus

966	12	0.9	196	11	08C2X4	08C2X4 mus musculus
967	12	0.9	239	5	025772	025772 plasmodium
968	12	0.9	246	5	09U4U4	09U4U4 cryptospori
969	12	0.9	263	5	09W2T4	09W2T4 drosophila
970	12	0.9	278	15	090UP8	090UP8 equine infe
971	12	0.9	301	5	08IE30	08IE30 plasmodium
972	12	0.9	306	5	09VM35	09VM35 drosophila
973	12	0.9	319	5	094499	094499 dictyosteli
974	12	0.9	325	5	094492	094492 dictyosteli
975	12	0.9	327	5	08MXP9	08MXP9 dictyosteli
976	12	0.9	334	5	08T2M7	08T2M7 dictyosteli
977	12	0.9	334	5	08I2B5	08I2B5 plasmodium
978	12	0.9	373	10	08L841	08L841 arabidopsis
979	12	0.9	387	3	08X0Z8	08X0Z8 candida alb
980	12	0.9	388	5	08TIM8	08TIM8 plasmodium
981	12	0.9	389	13	09PTK9	09PTK9 brachydanio
982	12	0.9	399	8	09B6E2	09B6E2 yarrowia 11
983	12	0.9	403	5	08IBV2	08IBV2 plasmodium
984	12	0.9	404	8	095946	095946 saccharomyc
985	12	0.9	405	10	093XP2	093XP2 zea mays (m
986	12	0.9	420	5	076853	076853 dictyosteli
987	12	0.9	427	5	09VI06	09VI06 drosophila
988	12	0.9	431	3	012124	012124 saccharomyc
989	12	0.9	451	5	095ZC1	095ZC1 leishmania
990	12	0.9	468	4	08NSM1	08NSM1 homo sapien
991	12	0.9	522	5	08MN27	08MN27 dictyosteli
992	12	0.9	551	5	08I1Y2	08I1Y2 plasmodium
993	12	0.9	563	5	08T6E1	08T6E1 dictyosteli
994	12	0.9	578	5	08I387	08I387 plasmodium
995	12	0.9	579	5	076734	076734 dictyosteli
996	12	0.9	588	5	08MMH8	08MMH8 dictyosteli
997	12	0.9	603	10	09C6K3	09C6K3 arabidopsis
998	12	0.9	611	5	08I449	08I449 plasmodium
999	12	0.9	616	5	08I5Q2	08I5Q2 plasmodium
1000	12	0.9	640	5	08I6Z8	08I6Z8 plasmodium

ALIGNMENTS

RESULT 1

```
Q814Y0 PRELIMINARY; PRT; 1351 AA.
AC Q814Y0:
ID Q814Y0:
AD Q814Y0:
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Kinesin-like protein, putative.
GN PFL2165W.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Hemosporidia; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368664;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.B., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chen M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Anguilo S.,
RA Petrea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McPadden G.I., Cummings L.M., Subramanian G.M., Mungai C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrett B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RT Nature 419:498-511(2002).
DR EMBL; AF014851; AAN36517.1; -.
SO SEQUENCE 1351 AA; 156007 MW; 8B69594BFD73C788 CRC64;
```

Query Match 100.0%; Score 1288; DB 5; Length 1351;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MSKIKYVAKRPLSELEKKKSDIITVKNCTIYIDERYKVDMTKIERHEFYDVK	60						
DB	64	MSKIKYVAKRPLSELEKKKSDIITVKNCTIYIDERYKVDMTKIERHEFYDVK	123						
QY	61	FDDTVNFTYENTIKFLIDLYENGVCSCFAYVGOTGSKTYMLGSPQYGSDPGIF	120						
DB	124	FDDTVNFTYENTIKFLIDLYENGVCSCFAYVGOTGSKTYMLGSPQYGSDPGIF	183						
QY	121	QYAAGDIPTFLNYDKDNTKGFISFEYIYCGKLYDLQKRVKVALENGKEVYVDLK	180						
DB	184	QYAAGDIPTFLNYDKDNTKGFISFEYIYCGKLYDLQKRVKVALENGKEVYVDLK	243						
QY	181	ILRLTKBELLMKIDGVLRLKIGVSNQDESSRAHLIYDKDKNTSLGKIAFDL	240						
DB	244	ILRLTKBELLMKIDGVLRLKIGVSNQDESSRAHLIYDKDKNTSLGKIAFDL	303						
QY	241	AGSERGADTVSONKQOTDGANINRSLALKECIRAMDSQKNI PFRDSELTKVLRDIFV	300						
DB	304	AGSERGADTVSONKQOTDGANINRSLALKECIRAMDSQKNI PFRDSELTKVLRDIFV	363						
QY	301	GKSKSIMIANISPTISCEBOTLTARYSVKQFKKSTCINBEDTNTERISILDSKGS	360						
DB	364	GKSKSIMIANISPTISCEBOTLTARYSVKQFKKSTCINBEDTNTERISILDSKGS	423						
QY	361	EMASSIENVYIKSNHLNKNKINRGKINIKIEBNNTLKQKSPDKREGFTSTGKY	420						
DB	424	EMASSIENVYIKSNHLNKNKINRGKINIKIEBNNTLKQKSPDKREGFTSTGKY	483						
QY	421	SSLNDIDIKKKKKGLINYKSTLYNDNTJNKGHNNNNNNNDNDNDNNNNNNNDSS	480						
DB	484	SSLNDIDIKKKKKGLINYKSTLYNDNTJNKGHNNNNNNNDNDNDNNNNNNNDSS	543						
QY	481	SMVNNMIMNINNNINNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	540						
DB	544	SMVNNMIMNINNNINNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	603						
QY	541	HLNNNDKSTFLHKKKRLBDNTKLKRRSSCDINMKKKKNTLHARRSVGSKLTMPFYDPOK	600						
DB	604	HLNNNDKSTFLHKKKRLBDNTKLKRRSSCDINMKKKKNTLHARRSVGSKLTMPFYDPOK	663						
QY	601	KDNFFKSNINKMEDNTPKDIYESRNVSNNGVLLGLNKNTHDISTDENNDKIN	660						
DB	664	KDNFFKSNINKMEDNTPKDIYESRNVSNNGVLLGLNKNTHDISTDENNDKIN	723						
QY	661	NGVINIINNSVNSINNSVNSINNSVNSINNSVNSINNSVNSINNSVNSINNSVNS	720						
DB	724	NGVINIINNSVNSINNSVNSINNSVNSINNSVNSINNSVNSINNSVNSINNSVNS	783						
QY	721	NNDIIFPDALSCDNNVPIITNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	780						
DB	784	NNDIIFPDALSCDNNVPIITNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	843						
QY	781	HNLFQPDNNKNTSNIQININTKKNQDGNVYVSNFCHYNLDKNYLLDLNKKQKDKNIH	840						
DB	844	HNLFQPDNNKNTSNIQININTKKNQDGNVYVSNFCHYNLDKNYLLDLNKKQKDKNIH	903						
QY	841	GCDNNIIONRDFEKKKKTTFYNNNNIYVYNNMGNNNSPRMKGLGSHTSIDNMKNE	900						
DB	904	GCDNNIIONRDFEKKKKTTFYNNNNIYVYNNMGNNNSPRMKGLGSHTSIDNMKNE	963						
QY	901	MKNMEMDKMKKHITSNNNNSSSSSSNNNIYNNINDDTFONDYCHNDNFTTIRKN	960						
DB	964	MKNMEMDKMKKHITSNNNNSSSSSSNNNIYNNINDDTFONDYCHNDNFTTIRKN	1023						
QY	961	NTNINSIYQNDIITYINSIDNYSNTLLHFKEKTYPTLSTNEDIYNNKMEKHAIRLD	1020						
DB	1024	NTNINSIYQNDIITYINSIDNYSNTLLHFKEKTYPTLSTNEDIYNNKMEKHAIRLD	1083						
QY	1021	DQKCYDDNNNNVNNKKNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNV	1080						
DB	1084	DQKCYDDNNNNVNNKKNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNV	1143						

Qy	1081	NFNNNNEYSYQKAVDVTIINNCLNSLDPSWYDPDKELINIIILSKYAEKONYKCCIN	1146
Db	1144	NFNNNNEYSYQKAVDVTIINNCLNSLDPSWYDPDKELINIIILSKYAEKONYKCCIN	1203
Qy	1141	EDIKNMSLEEDIKTAOSIYERKRKVLTYKLLLFKKNAVDTQINNETSDLRKOLVWCHICCN	1200
Db	1204	EDIKNMSLEEDIKTAOSIYERKRKVLTYKLLLFKKNAVDTQINNETSDLRKOLVWCHICCN	1263
Qy	1201	NPDDQFHFYASRLKEDIINLIMRQIWCESENRLYQFLVVEYQNSKANSVLINVSSN	1260
Db	1264	NPDDQFHFYASRLKEDIINLIMRQIWCESENRLYQFLVVEYQNSKANSVLINVSSN	1323
Qy	1261	NGDIITILNKKKLVODNIKSMOHNNTIHKK	1288
Db	1324	NGDIITILNKKKLVODNIKSMOHNNTIHKK	1351

RESULT 2	08T224	PRELIMINARY;	PRT; 1723 AA.
AC	08T224		
DT	01-JUN-2002 (TrEMBLrel. 21, Created)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Hypothetical 188.9 kDa protein.		
OS	Dictyostelium discoidium (Slime mold).		
OC	Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.		
OX	NCBI_TaxId=44689;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=AX4;		
RA	Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,		
RA	Leimann R., Baumgart C., Parra G., April J.F., Guiso R., Kumpf K.,		
RA	Tunggal B., Cox E., Quail M.A., Platzner M., Rosenthal A., Noegel A.A.;		
RL	"Sequence and Analysis of Chromosome 2 of Dictyostelium";		
RL	Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.		
DR	EMBL; AC116030; AL92981.1; -		
DR	InterPro; IPR002048; EF-hand.		
DR	InterPro; IPR001356; Homeobox.		
DR	Pfam; PF00036; efhand.1.		
DR	Pfam; PF00046; homeobox.1.		
DR	SMART; SM00389; HOX.1.		
DR	PROSITE; PSS0071; HOMEBOX_2.1.		
KW	Hypothetical protein.		
SO	SEQUENCE 1723 AA; 188915 MW; 686A3B702B3FEB95 CRC64;		

```

Query Match          1.5%; Score 19; DB 5; Length 1723;
Best Local Similarity 100.0%; Pred. No. 6,1e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      739 NITNNNNNNNNNNNNNNNNNNN 757
        |||
Db       544 NITNNNNNNNNNNNNNNNNNNN 562

RESULT 3
QB8MN71 ID QB8MN71 PRELIMINARY; PRT; 348 AA.
AC QB8MN71;
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tungal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
```

RT "Sequence and Analysis of Chromosome 2 of Dictyostelium."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC117079; AM43722.1; -.
KW Hypothetical protein.
SQ SEQUENCE 348 AA; 39051 MW; 9A3DDDC4941FF548 CRC64;

Query Match 1.4%; Score 18; DB 5; Length 348;
Best Local Similarity 100.0%; Pred. No. 1.5e+08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	742	NNNNNNNNNNNNNNNNNNID	759
Db	164	NNNNNNNNNNNNNNNNNNID	181

RESULT 4	
Q6T832	PRELIMINARY; PRT; 439 AA.
ID Q6T832	
AC Q6T832;	
DT 01-JUN-2002 (TrEMBLrel. 21, Created)	
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)	
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)	
DE Hypothetical 51.5 kDa protein.	
OS Dictyostelium discoideum (Stime mold).	
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.	
OX NCBI_TaxID=44689;	
LN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=AX4;	
RA Gloeckner G., Eichinger L., Szafranski K., Pachebt J., Dear P., Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K., Rattmann B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A., "Sequence and Analysis of Chromosome 2 of Dictyostelium "; Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.	
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.	
DR EMBL; ACL15581; AAL92229.1; -.	
DR Hypothetical protein.	
Q6 SEQUENCE 439 AA; 51503 MW; D8BC6EAD046B0C2E CRC64;	

```
Query Match      1.4%; Score 18; DB 5; Length 439;  
Best Local Similarity 100.0%; Pred. No. 1 Be-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 5
Q8MN10 PRELIMINARY; PRT; 699 AA
Q8MN10

DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxId=44689;
RN [1]
RP
RC
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tungal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium.";
RL Submitted (MAY-2002) to the EMBL/genbank/DBJ databases.
DR EMBL; ACU11717; AAA43662.1; --
KW Hypothetical protein.
SQ
SQ SEQUENCE 699 AA; 79820 MW; 5A1D3995C439AEAD CRC64;

Query Match	1.4%;	Score 18;	DB 5;	Length 699;
Best Local Similarity	100.0%;	Pred. No. 2.7e-08;		
Matches	18;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;

RESULT 13

Q8MP24 ID Q8MP24 PRELIMINARY; PRT; 172 AA.
AC Q8MP24;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC123513; AAM44372.1; -.
KW Hypothetical protein.
SQ SEQUENCE 172 AA; 20800 MW; CE6A456B2B3CFDCE CRC64;

Query Match 1.3%; Score 17; DB 5; Length 172;
Best Local Similarity 100.0%; Pred. No. 7.9e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 741 TNNNNNNNNNNNNNNNNNN 757
Db 111 TNNNNNNNNNNNNNNNNNN 127

RESULT 14

Q8ST26 ID Q8ST26 PRELIMINARY; PRT; 204 AA.
AC Q8ST26;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE HSPC305 10/100.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC115575; AAL92188.1; -.
SQ SEQUENCE 204 AA; 23921 MW; EFCD3733C7BA9C81 CRC64;

Query Match 1.3%; Score 17; DB 5; Length 204;
Best Local Similarity 100.0%; Pred. No. 9.1e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 742 NNNNNNNNNNNNNNNNNNI 758
Db 156 NNNNNNNNNNNNNNNNNNI 172

RESULT 15

Q8T1B5 ID Q8T1B5 PRELIMINARY; PRT; 279 AA.
AC Q8T1B5;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical 31.4 kDa protein.

OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116960; AAM08485.1; -.
KW Hypothetical protein.
SQ SEQUENCE 279 AA; 31413 MW; 182A4849594CBD04 CRC64;

Query Match 1.3%; Score 17; DB 5; Length 279;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 741 TNNNNNNNNNNNNNNNNNN 757
Db 100 TNNNNNNNNNNNNNNNNNN 116

RESULT 16

Q8ST23 ID Q8ST23 PRELIMINARY; PRT; 284 AA.
AC Q8ST23;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE HDCKB03P 10/100.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC115681; AAL92673.1; -.
SQ SEQUENCE 284 AA; 32029 MW; B52D22929F7D526A CRC64;

Query Match 1.3%; Score 17; DB 5; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 742 NNNNNNNNNNNNNNNNNNI 758
Db 197 NNNNNNNNNNNNNNNNNNI 213

RESULT 17

Q9XY02 ID Q9XY02 PRELIMINARY; PRT; 288 AA.
AC Q9XY02;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE PLOX5-DJ (Fragment).
GN PLOX5-DJ.
OS Dugesia japonica (Planarian).
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;
OC Paludicola; Dugesitidae; Dugesia.
OX NCBI_TaxID=6161;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GI;
RX MEDLINE=99287788; PubMed=10357903;
RA Orit H., Kato K., Umesono Y., Sakurai T., Agata K., Watanabe K.;

RT "The planarian Hox/Hox homeobox genes (Plox) expressed along the
anterior-posterior axis." ;
RL Dev. Biol. 210:456-468(1999).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: AB024409; BAA7405.1; -.
DR HSSP: P02833; 9ANT.
DR InterPro: IPR001827; Antennapedia.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR PRINTS: PR00025; ANTENNAPEDIA.
DR PRINTS: PR00024; HOMEBOX.
DR ProDom: PD000010; Homeobox; 1.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.
KM DNA-binding; Homeobox; Nuclear protein.
FT NON TER 1
SQ SEQUENCE 288 AA; 33191 MW; 693C649656B290A CRC64;

Query Match 1.3%; Score 17; DB 5; Length 288;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 741 TNNNNNNNNNNNNNNNN 757
Db 146 TNNNNNNNNNNNNNNNN 162

RESULT 18
Q8MX75 PRELIMINARY; PRT; 293 AA.
ID Q8MX75;
AC Q8MX75;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CP47 (Fragment).
GN CP47.
OS Cryosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryosporidiidae; Cryosporidium.
OC NCBI_TaxID=5807;
RN [1]
RP SEQUENCE FROM N.A.
RA Jenkins M.C., Trout J.M., Fayer R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF384127; AAM46174.1; -.
FT NON TER 1
SQ SEQUENCE 293 AA; 32688 MW; 7C81729D50B46E0D CRC64;

Query Match 1.3%; Score 17; DB 5; Length 293;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 741 TNNNNNNNNNNNNNNNN 757
Db 97 TNNNNNNNNNNNNNNNN 113

RESULT 19
Q8MNN2 PRELIMINARY; PRT; 300 AA.
ID Q8MNN2;
AC Q8MNN2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidida; Dictyostelium.
OC NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=KX4;
RC Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,

RA Lehmann R., Baumgart C., Parra G., April J.F., Guiso R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC116965; AAM33218.1; -.
DR InterPro: IPR001159; DS_RBD.
DR Pfam: PF00035; dsrm; 1.
DR SMART: SM00358; DSRM; 1.
DR PROSITE: PS50137; DS_RBD; 1.
KM Hypothetical protein.
SQ SEQUENCE 300 AA; 35188 MW; C4EC727AF6C0DE2D CRC64;

Query Match 1.3%; Score 17; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNN 758
Db 185 NNNNNNNNNNNNNNNNN 201

RESULT 20
Q8ICP1 PRELIMINARY; PRT; 371 AA.
ID Q8ICP1;
AC Q8ICP1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN MAL6P1.298.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OC NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Cherevach I., Davis P., Goodhead I., Stevens K., Mungall K.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrett B.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL844505; CND50400.1; -.
KM Hypothetical protein.
SQ SEQUENCE 371 AA; 43460 MW; 8A81197CA8F81BC2 CRC64;

Query Match 1.3%; Score 17; DB 5; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNN 758
Db 227 NNNNNNNNNNNNNNNNN 243

RESULT 21
Q8IHR1 PRELIMINARY; PRT; 384 AA.
ID Q8IHR1;
AC Q8IHR1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN PF11.0468.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OC NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=2255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Valdiva A.B.,


```

RA Martin D.M.A., Fairclamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McPadden G.I., Cummings L.M., Subramanian G.M., Murgall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrett B.,
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
DR EMBL: AE014843; AAM36048.1; -
KW Hypothetical protein.
SQ SEQUENCE 384 AA; 46253 MW; F20DP4F1EF5089CB CRC64;

Query Match 1.3%; Score 17; DB 5; Length 384;
Best Local Similarity 100.0%; Pred.No. 1.5e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 741 TNNNNNNNNNNNNNNNNNN 757
Db 211 TNNNNNNNNNNNNNNNNNN 227

RESULT 22
OY 025768 PRELIMINARY; PRT; 391 AA.
ID 025768;
AC 025768;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Asparagine-rich antigen (Clone 14C1) (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Schreiber W.G., Deutsche U., Storck T., Mueller-Hill B.;
RL Submitted (DEC-1989) to the EMBL/GenBank/DBJ databases.
DR EMBL: X17483; CAA35519.1; -.
FT NON TER 1
FT NON TER 391
SQ SEQUENCE 391 AA; 45309 MW; 336596F24F071A8 CRC64;

Query Match 1.3%; Score 17; DB 5; Length 391;
Best Local Similarity 100.0%; Pred.No. 1.6e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 741 TNNNNNNNNNNNNNNNNNN 757
Db 162 TNNNNNNNNNNNNNNNNNN 178

RESULT 23
OY 025774 PRELIMINARY; PRT; 419 AA.
ID 025774;
AC 025774;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Asparagine-rich antigen (Clone 52C11) (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Schreiber W.G., Deutsche U., Storck T., Mueller-Hill B.;
RL Submitted (DEC-1989) to the EMBL/GenBank/DBJ databases.
DR EMBL: X17489; CAA35525.1; -.
FT NON TER 1
FT NON TER 419
SQ SEQUENCE 419 AA; 47365 MW; 08DFBD456298DE31 CRC64;

Query Match 1.3%; Score 17; DB 5; Length 419;
Best Local Similarity 100.0%; Pred.No. 1.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 741 TNNNNNNNNNNNNNNNNNN 757

```

```

Db 25 TNNNNNNNNNNNNNNNNNN 41

RESULT 24
OY 08MNF6 PRELIMINARY; PRT; 426 AA.
ID 08MNF6;
AC 08MNF6;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC117072; AAM33165.1; -.
KW Hypothetical protein.
SQ SEQUENCE 426 AA; 48800 MW; 10D36C7BA3AC69A CRC64;

Query Match 1.3%; Score 17; DB 5; Length 426;
Best Local Similarity 100.0%; Pred.No. 1.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 742 NNNNNNNNNNNNNNNNNNNI 758
Db 401 NNNNNNNNNNNNNNNNNNNI 417

RESULT 25
OY 08T834 PRELIMINARY; PRT; 434 AA.
ID 08T834;
AC 08T834;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical 49.6 kDa protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC115581; AAL92226.1; -.
DR InterPro: IPR001594; Znf_DHHC.
DR Pfam: PF01528; Zf-DHHC_1.
DR ProDom: PD003041; Znf_DHHC_1.
DR PROSITE: PS50216; ZF_DHHC; 1.
KW Hypothetical protein.
SQ SEQUENCE 434 AA; 49641 MW; 256228913AAS5E9A CRC64;

Query Match 1.3%; Score 17; DB 5; Length 434;
Best Local Similarity 100.0%; Pred.No. 1.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 742 NNNNNNNNNNNNNNNNNNNI 758
Db 288 NNNNNNNNNNNNNNNNNNNI 304

RESULT 26

```



```

0814W5 ID 0814W5 PRELIMINARY; PRT; 440 AA.
AC 0814W5;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
GN Hypothetical protein.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=2255705; PubMed=1236864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,
RA Perte M., Allen J., Selengut J., Haft D., Mather M.W., Valdivia A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511 (2002).
DR EMBL; AE014851; AAN36532.1; -.
KW Hypothetical protein.
SQ SEQUENCE 440 AA; 51713 MW; E2FBF15E1039566 CRC64;

Query Match 1.3%; Score 17; DB 5; Length 440;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 741 TNNNNNNNNNNNNNNNNNNNN 757
DB 253 TNNNNNNNNNNNNNNNNNNNN 269

RESULT 27
Q8MXW8 ID 08MXW8 PRELIMINARY; PRT; 455 AA.
AC 08MXW8;
DT 01-OCT-2002 (TEMBLrel. 22, Created)
DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
GN Winged helix/forkhead transcription factor (Fragment).
OS Dugesia japonica (Planarian).
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Sericata; Tricladida;
OC Paludicola; Dugesidae; Dugesia.
OX NCBI_TaxID=6161;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GI;
RA Kojima S., Umesono Y., Matanabe K., Agata K.;
RT "The expression of planarian brain factor homologues, DfFoxg and
RT DfFoxD."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB091062; BAC10918.1; -.
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00250; Fork head; 1.
DR PRINTS; PRO0053; FORKHEAD.
DR PRODOM; PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS50039; FORK_HEAD_3; 1.
FT NON TER 1
SQ SEQUENCE 455 AA; 53245 MW; 1F35B9F5BEEF2140 CRC64;

Query Match 1.3%; Score 17; DB 5; Length 455;

```

```

Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 741 TNNNNNNNNNNNNNNNNNNNN 757
DB 357 TNNNNNNNNNNNNNNNNNNNN 373

RESULT 28
Q81JG6 ID 081JG6 PRELIMINARY; PRT; 465 AA.
AC 081JG6;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
GN Ubiquitin C-terminal hydrolase, family 1, putative.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=2255705; PubMed=1236864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,
RA Perte M., Allen J., Selengut J., Haft D., Mather M.W., Valdivia A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511 (2002).
DR EMBL; AE014838; AAN35761.1; -.
KW Hydrolase.
SQ SEQUENCE 465 AA; 54460 MW; 138A55EB774F99DD CRC64;

Query Match 1.3%; Score 17; DB 5; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 TNNNNNNNNNNNNNNNNNNNN 758
DB 277 TNNNNNNNNNNNNNNNNNNNN 293

RESULT 29
Q8SSX7 ID 08SSX7 PRELIMINARY; PRT; 496 AA.
AC 08SSX7;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
GN Hypothetical protein B01A2.2_9/101.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachbat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tungal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC115685; AML92706.1; -.
DR InterPro; IPR00504; RNA_rec_mot.
DR PROSITE; PS50102; RRM; 1.
SQ SEQUENCE 496 AA; 56227 MW; 614DC0EDED769126 CRC64;

```


Query Match 1.3%; Score 17; DB 5; Length 496;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNNNNI 758
DB 94 NNNNNNNNNNNNNNNNNNNI 110

RESULT 30

Q8MP37 PRELIMINARY; PRT; 498 AA.
ID Q8MP37;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelid; Dictyostelium.
CX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC117078; AAM45290.1; -.
DR InterPro; IPR003034; SAP.
KW Hypothetical protein.
SQ SEQUENCE 498 AA; 57878 MW; ADA58F246919E9A CRC64;

Query Match 1.3%; Score 17; DB 5; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNNNNI 758
DB 464 NNNNNNNNNNNNNNNNNNNI 480

RESULT 31

Q8SSZ5 PRELIMINARY; PRT; 511 AA.
ID Q8SSZ5;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical protein. 6/101.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelid; Dictyostelium.
CX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116330; AAL96721.1; -.
DR InterPro; IPR003734; DUF155.
DR Pfam; PF02582; DUF155; 1.
KW Hypothetical protein.
SQ SEQUENCE 511 AA; 58641 MW; DLF00B910D2B0D5 CRC64;

Query Match 1.3%; Score 17; DB 5; Length 511;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNNNNI 758
DB 464 NNNNNNNNNNNNNNNNNNNI 480

DB 178 NNNNNNNNNNNNNNNNNNNI 194

RESULT 32

Q8T1K7 PRELIMINARY; PRT; 537 AA.
ID Q8T1K7;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical 63.3 kDa protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelid; Dictyostelium.
CX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116425; AAL99099.1; -.
KW Hypothetical protein.
SQ SEQUENCE 537 AA; 63290 MW; A935AD400FE8137D CRC64;

Query Match 1.3%; Score 17; DB 5; Length 537;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNNNNI 758
DB 244 NNNNNNNNNNNNNNNNNNNI 260

RESULT 33

Q61085 PRELIMINARY; PRT; 560 AA.
ID Q61085;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Coronin binding protein.
GN DB10.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelid; Dictyostelium.
CX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RA de Hostos E.L., Benhayon D., Berger M., Gu W.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF051898; AAC05577.1; -.
SQ SEQUENCE 560 AA; 61281 MW; P417511014157229 CRC64;

Query Match 1.3%; Score 17; DB 5; Length 560;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 741 TNNNNNNNNNNNNNNNNNNNN 757
DB 454 TNNNNNNNNNNNNNNNNNNNN 470

RESULT 34

Q8MNG3 PRELIMINARY; PRT; 587 AA.
ID Q8MNG3;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelid; Dictyostelium.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: October 2, 2003, 16:35:32 ; Search time 85 Seconds
(without alignments)
2397.392 Million cell updates/sec

Title: US-10-006-780-2
Perfect score: 1288
Sequence: 1 NMSKIKVVRKRPRLSELEKK.....KKLVODINKSMDDNNHHKK 1288

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 587654 seqs, 158212981 residues

Word size : 0

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

1: Published Applications_AAI*
2: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1288	100.0	1288	US-10-006-780-2	Sequence 2, Appli
2	354	27.5	355	US-10-006-780-6	Sequence 6, Appli
3	332	25.8	332	US-10-006-780-4	Sequence 4, Appli
4	332	25.8	332	US-10-006-780-10	Sequence 10, Appli
5	330	25.6	361	US-10-006-780-8	Sequence 8, Appli
6	17	1.3	666	US-09-801-368-36	Sequence 36, Appli
7	16	1.2	213	US-10-032-585-7714	Sequence 7714, Ap
8	16	1.2	732	US-09-994-485-8	Sequence 8, Appli
9	16	1.2	732	US-09-832-292-12	Sequence 12, Appli
10	16	1.2	758	US-09-801-368-224	Sequence 224, App
11	16	1.2	798	US-10-033-558-7565	Sequence 7565, Ap
12	16	1.2	961	US-09-801-368-132	Sequence 132, App
13	16	1.2	1331	US-10-087-464-50	Sequence 50, Appli
14	16	1.2	2150	US-10-135-322-17	Sequence 17, Appli
15	16	1.2	2184	US-10-304-095-6	Sequence 6, Appli

16	1.2	2690	15	US-10-184-644-35	Sequence 35, Appli
17	1.2	2690	15	US-10-184-644-35	Sequence 35, Appli
18	1.2	73	9	US-09-864-761-47535	Sequence 47535, A
19	1.2	303	12	US-10-422-866-4	Sequence 4, Appli
20	1.2	492	12	US-10-422-866-2	Sequence 2, Appli
21	1.2	794	12	US-10-304-095-8	Sequence 8, Appli
22	1.2	888	9	US-09-826-752-6	Sequence 6, Appli
23	1.1	375	9	US-09-839-185-2	Sequence 2, Appli
24	1.1	458	12	US-10-033-585-7575	Sequence 7575, Ap
25	1.1	1662	15	US-10-195-144-37	Sequence 37, Appli
26	1.1	1770	10	US-09-801-368-288	Sequence 298, App
27	1.0	394	12	US-10-032-585-7237	Sequence 7237, Ap
28	1.0	941	12	US-10-032-585-7930	Sequence 7930, Ap
29	1.0	1151	10	US-09-801-368-134	Sequence 134, App
30	1.0	1151	15	US-10-060-019-31	Sequence 31, Appli
31	0.9	387	12	US-10-032-585-7034	Sequence 7034, Ap
32	0.9	410	10	US-09-925-300-1074	Sequence 1074, Ap
33	0.9	511	12	US-10-032-585-7788	Sequence 7788, Ap
34	0.9	725	12	US-09-849-602-20	Sequence 20, Appli
35	0.9	1501	10	US-09-924-154-17	Sequence 17, Appli
36	0.9	160	15	US-10-106-698-4938	Sequence 4938, Ap
37	0.9	250	11	US-09-934-455-462	Sequence 462, App
38	0.9	298	15	US-10-286-264-36	Sequence 36, Appli
39	0.9	298	16	US-10-278-536-188	Sequence 188, App
40	0.9	309	10	US-09-991-496-134	Sequence 134, App
41	0.9	314	11	US-09-893-519A-30	Sequence 30, Appli
42	0.9	319	10	US-09-723-276-2	Sequence 2, Appli
43	0.9	338	10	US-09-883-096-4	Sequence 4, Appli
44	0.9	343	15	US-10-159-151-8	Sequence 8, Appli
45	0.9	363	15	US-10-159-151-10	Sequence 10, Appli
46	0.9	373	15	US-10-159-151-4	Sequence 4, Appli
47	0.9	381	15	US-10-159-151-6	Sequence 6, Appli
48	0.9	383	10	US-09-883-096-5	Sequence 5, Appli
49	0.9	409	12	US-10-032-585-7054	Sequence 7054, Ap
50	0.9	676	10	US-09-801-368-302	Sequence 302, App
51	0.9	864	10	US-09-883-096-2	Sequence 2, Appli
52	0.9	1230	12	US-10-093-524-6	Sequence 6, Appli
53	0.9	1369	12	US-10-342-224-62	Sequence 62, Appli
54	0.9	1368	15	US-10-159-151-2	Sequence 2, Appli
55	0.9	1388	15	US-10-146-473-82	Sequence 82, Appli
56	0.8	16	11	US-09-533-382-40	Sequence 40, Appli
57	0.8	16	15	US-10-067-790-40	Sequence 40, Appli
58	0.8	16	15	US-10-067-892-40	Sequence 40, Appli
59	0.8	16	15	US-10-067-893-40	Sequence 40, Appli
60	0.8	40	12	US-10-083-768-226	Sequence 226, App
61	0.8	265	12	US-10-032-585-7166	Sequence 7166, Ap
62	0.8	412	10	US-09-788-345-12	Sequence 12, Appli
63	0.8	412	12	US-10-337-312-12	Sequence 12, Appli
64	0.8	488	12	US-10-033-585-7710	Sequence 7710, Ap
65	0.8	500	12	US-10-320-646-39	Sequence 29, Appli
66	0.8	500	12	US-10-320-646-34	Sequence 34, Appli
67	0.8	501	12	US-10-320-646-37	Sequence 27, Appli
68	0.8	501	12	US-10-320-646-28	Sequence 28, Appli
69	0.8	511	9	US-09-949-780-2	Sequence 2, Appli
70	0.8	522	9	US-09-965-313-7	Sequence 7, Appli
71	0.8	523	12	US-10-017-161-1982	Sequence 1982, Ap
72	0.8	544	12	US-10-032-585-7849	Sequence 7849, Ap
73	0.8	596	12	US-10-167-831-2	Sequence 2, Appli
74	0.8	623	14	US-10-108-605-125	Sequence 125, App
75	0.8	623	14	US-10-108-605-129	Sequence 129, App
76	0.8	659	10	US-09-833-555-20	Sequence 20, Appli
77	0.8	690	9	US-09-334-477-47	Sequence 47, Appli
78	0.8	694	9	US-09-334-477-49	Sequence 49, Appli
79	0.8	708	9	US-09-334-477-33	Sequence 33, Appli
80	0.8	711	9	US-09-334-477-35	Sequence 35, Appli
81	0.8	824	10	US-09-801-368-112	Sequence 312, App
82	0.8	861	11	US-09-820-843A-109	Sequence 242, App
83	0.8	1068	10	US-09-801-368-242	Sequence 242, App
84	0.8	1420	12	US-10-241-586-116	Sequence 116, App
85	0.8	1598	12	US-10-032-585-7766	Sequence 7766, Ap
86	0.8	1648	15	US-10-157-031-38	Sequence 38, Appli
87	0.7	34	9	US-09-864-761-48088	Sequence 48088, A
88	0.7	57	9	US-09-864-761-43371	Sequence 43371, A

89	9	0.7	177	12	US-09-903-190-164	Sequence 164, App	162	8	0.6	1015	12	US-10-032-585-7162	Sequence 7162, App
90	9	0.7	354	12	US-10-032-585-7079	Sequence 7079, App	163	8	0.6	1115	12	US-10-335-711-71	Sequence 2, Appl1
91	9	0.7	440	15	US-10-083-357-1256	Sequence 1256, App	164	8	0.6	1207	14	US-10-108-605-71	Sequence 71, Appl1
92	9	0.7	450	12	US-10-032-585-7816	Sequence 7816, App	165	8	0.6	1242	10	US-09-903-248-5	Sequence 5, Appl1
93	9	0.7	498	12	US-10-032-585-7230	Sequence 7230, App	166	8	0.6	1242	10	US-09-859-664-5	Sequence 5, Appl1
94	9	0.7	564	10	US-09-801-368-190	Sequence 190, App	167	8	0.6	1242	10	US-09-903-063-5	Sequence 5, Appl1
95	9	0.7	670	12	US-10-032-585-7568	Sequence 7568, App	168	8	0.6	1242	10	US-09-903-216-5	Sequence 5, Appl1
96	9	0.7	715	12	US-10-171-404-44	Sequence 44, App	169	8	0.6	1242	10	US-09-903-023-5	Sequence 5, Appl1
97	9	0.7	740	12	US-10-032-585-7673	Sequence 7673, App	170	8	0.6	1242	10	US-09-903-023-5	Sequence 5, Appl1
98	9	0.7	769	12	US-10-032-585-7597	Sequence 7597, App	171	8	0.6	1242	11	US-09-436-184-5	Sequence 5, Appl1
99	9	0.7	769	12	US-10-032-585-7117	Sequence 7117, App	172	8	0.6	1242	14	US-10-085-027-1	Sequence 11, Appl1
100	9	0.7	807	11	US-09-820-843A-108	Sequence 108, App	173	8	0.6	1284	9	US-09-119-855-11	Sequence 9, Appl1
101	9	0.7	856	12	US-09-949-029-66	Sequence 66, App	174	8	0.6	1284	15	US-10-160-224-9	Sequence 11, Appl1
102	9	0.7	887	12	US-09-949-029-130	Sequence 130, App	175	8	0.6	1403	14	US-10-108-605-93	Sequence 93, Appl1
103	9	0.7	928	10	US-09-801-368-282	Sequence 282, App	176	8	0.6	1403	14	US-10-108-605-93	Sequence 5, Appl1
104	9	0.7	1143	12	US-10-032-585-7789	Sequence 7789, App	177	8	0.6	1850	12	US-10-146-657-5	Sequence 5, Appl1
105	9	0.7	1165	14	US-10-016-768-10	Sequence 10, App	178	7	0.5	10	12	US-09-572-270A-530	Sequence 530, App
106	9	0.7	1207	12	US-09-949-029-144	Sequence 144, App	179	7	0.5	10	12	US-09-572-270A-532	Sequence 532, App
107	9	0.7	1221	14	US-10-016-768-11	Sequence 11, App	180	7	0.5	10	12	US-09-572-270A-534	Sequence 534, App
108	9	0.7	1245	12	US-10-032-585-7657	Sequence 7657, App	181	7	0.5	10	12	US-09-572-270A-536	Sequence 536, App
109	9	0.7	3092	10	US-09-801-368-172	Sequence 172, App	182	7	0.5	10	12	US-09-572-270A-538	Sequence 538, App
110	8	0.6	10	12	US-09-572-270A-565	Sequence 565, App	183	7	0.5	10	12	US-09-572-270A-544	Sequence 544, App
111	8	0.6	39	9	US-09-864-761-41774	Sequence 41774, A	184	7	0.5	10	12	US-09-572-270A-546	Sequence 546, App
112	8	0.6	41	9	US-09-864-761-4443	Sequence 4443, A	185	7	0.5	10	12	US-09-572-270A-548	Sequence 548, App
113	8	0.6	52	9	US-09-864-761-33973	Sequence 33973, A	186	7	0.5	10	12	US-09-572-270A-550	Sequence 550, App
114	8	0.6	75	9	US-09-864-761-44431	Sequence 44431, A	187	7	0.5	10	12	US-09-572-270A-552	Sequence 552, App
115	8	0.6	90	15	US-10-253-007-59	Sequence 59, App	188	7	0.5	10	12	US-09-572-270A-554	Sequence 554, App
116	8	0.6	146	15	US-10-180-375-138	Sequence 138, App	189	7	0.5	10	12	US-09-572-270A-567	Sequence 567, App
117	8	0.6	164	9	US-09-745-763-104	Sequence 104, App	190	7	0.5	10	12	US-09-572-270A-569	Sequence 569, App
118	8	0.6	175	14	US-10-108-605-97	Sequence 97, Appl1	191	7	0.5	10	12	US-09-572-270A-834	Sequence 834, App
119	8	0.6	195	12	US-10-314-372-4	Sequence 4, Appl1	192	7	0.5	10	12	US-09-572-270A-838	Sequence 838, App
120	8	0.6	221	15	US-10-021-811-14	Sequence 14, Appl1	193	7	0.5	10	12	US-09-572-270A-1008	Sequence 1008, App
121	8	0.6	222	12	US-10-032-585-7349	Sequence 7349, App	194	7	0.5	10	12	US-09-572-270A-1010	Sequence 1010, App
122	8	0.6	229	9	US-09-925-301-1283	Sequence 1283, App	195	7	0.5	16	9	US-09-935-682-11	Sequence 11, Appl1
123	8	0.6	234	8	US-08-808-031A-35	Sequence 35, Appl1	196	7	0.5	17	9	US-09-864-761-42401	Sequence 42401, A
124	8	0.6	250	11	US-09-934-455-256	Sequence 256, App	197	7	0.5	17	11	US-09-876-904A-233	Sequence 233, App
125	8	0.6	250	15	US-10-180-375-194	Sequence 194, App	198	7	0.5	19	9	US-09-864-761-37256	Sequence 7256, A
126	8	0.6	251	12	US-10-032-585-7420	Sequence 7420, App	199	7	0.5	22	9	US-09-864-761-40447	Sequence 40447, A
127	8	0.6	251	15	US-10-295-403-142	Sequence 142, App	200	7	0.5	23	11	US-09-967-869A-1	Sequence 1, Appl1
128	8	0.6	256	9	US-09-925-301-1083	Sequence 1083, App	201	7	0.5	24	14	US-10-123-427-7	Sequence 7, Appl1
129	8	0.6	259	8	US-08-808-031A-5	Sequence 5, Appl1	202	7	0.5	25	9	US-09-864-761-45219	Sequence 45219, A
130	8	0.6	300	9	US-09-861-451A-42	Sequence 42, Appl1	203	7	0.5	26	9	US-09-864-761-44765	Sequence 44765, A
131	8	0.6	302	9	US-09-443-704-12	Sequence 12, Appl1	204	7	0.5	30	9	US-09-864-761-41595	Sequence 41595, A
132	8	0.6	302	14	US-10-008-118A-12	Sequence 12, Appl1	205	7	0.5	30	9	US-09-864-761-42325	Sequence 42325, A
133	8	0.6	312	11	US-09-934-455-160	Sequence 160, App	206	7	0.5	31	10	US-09-764-846-248	Sequence 248, App
134	8	0.6	324	9	US-09-874-923-53	Sequence 53, Appl1	207	7	0.5	31	15	US-10-091-483-248	Sequence 248, App
135	8	0.6	324	10	US-09-991-496-53	Sequence 53, Appl1	208	7	0.5	32	9	US-09-864-761-36039	Sequence 36039, A
136	8	0.6	333	15	US-10-278-173-146	Sequence 146, App	209	7	0.5	33	9	US-09-864-761-3613	Sequence 3613, A
137	8	0.6	341	12	US-10-017-161-2038	Sequence 2038, App	210	7	0.5	33	9	US-09-864-761-47462	Sequence 47462, A
138	8	0.6	342	15	US-10-132-652-3	Sequence 3, Appl1	211	7	0.5	43	14	US-10-124-557-12	Sequence 12, Appl1
139	8	0.6	354	10	US-09-801-368-414	Sequence 414, App	212	7	0.5	44	9	US-09-864-761-43901	Sequence 43901, A
140	8	0.6	372	15	US-10-132-652-2	Sequence 2, Appl1	213	7	0.5	46	9	US-09-864-761-45070	Sequence 45070, A
141	8	0.6	376	15	US-10-278-173-142	Sequence 142, App	214	7	0.5	51	12	US-10-008-524A-100	Sequence 100, App
142	8	0.6	408	9	US-09-802-213-9	Sequence 9, Appl1	215	7	0.5	52	9	US-09-864-761-43647	Sequence 43647, App
143	8	0.6	419	11	US-09-934-455-422	Sequence 422, App	216	7	0.5	54	9	US-09-864-761-34566	Sequence 34566, A
144	8	0.6	420	12	US-10-032-585-7653	Sequence 7653, App	217	7	0.5	58	9	US-09-726-643-59	Sequence 59, Appl1
145	8	0.6	430	12	US-09-769-736-6	Sequence 6, Appl1	218	7	0.5	58	14	US-10-042-141-59	Sequence 59, Appl1
146	8	0.6	464	12	US-10-161-051-119	Sequence 119, App	219	7	0.5	63	9	US-09-877-258-2	Sequence 2, Appl1
147	8	0.6	528	12	US-10-032-585-7571	Sequence 7571, App	220	7	0.5	63	9	US-09-864-761-42311	Sequence 42311, A
148	8	0.6	554	10	US-09-801-368-142	Sequence 142, App	221	7	0.5	63	10	US-09-764-846-260	Sequence 260, App
149	8	0.6	555	8	US-08-808-031A-11	Sequence 31, Appl1	222	7	0.5	63	15	US-10-091-483-260	Sequence 260, App
150	8	0.6	586	8	US-08-808-031A-28	Sequence 28, Appl1	223	7	0.5	66	10	US-09-764-877-1302	Sequence 1302, App
151	8	0.6	605	10	US-09-758-269-16	Sequence 16, Appl1	224	7	0.5	71	10	US-09-738-626-6442	Sequence 6442, App
152	8	0.6	652	9	US-09-351-794A-2	Sequence 2, Appl1	225	7	0.5	72	9	US-09-764-869-683	Sequence 683, App
153	8	0.6	653	9	US-09-746-801A-13	Sequence 13, Appl1	226	7	0.5	72	15	US-10-091-504-683	Sequence 683, App
154	8	0.6	710	12	US-10-032-585-7813	Sequence 7813, App	227	7	0.5	74	15	US-10-178-213-368	Sequence 368, App
155	8	0.6	721	11	US-10-286-264-70	Sequence 70, Appl1	228	7	0.5	94	10	US-09-764-864-1273	Sequence 1273, App
156	8	0.6	725	11	US-09-934-455-442	Sequence 442, App	229	7	0.5	100	9	US-09-922-138-70	Sequence 20, Appl1
157	8	0.6	770	9	US-09-303-232-2	Sequence 2, Appl1	230	7	0.5	104	9	US-09-858-664A-10	Sequence 10, Appl1
158	8	0.6	809	12	US-10-032-585-7691	Sequence 7691, App	231	7	0.5	104	9	US-09-864-761-41552	Sequence 41552, A
159	8	0.6	834	9	US-09-826-752-4	Sequence 4, Appl1	232	7	0.5	104	10	US-09-764-864-1276	Sequence 1276, App
160	8	0.6	854	12	US-10-335-711-9	Sequence 9, Appl1	233	7	0.5	105	14	US-10-135-322-6	Sequence 6, Appl1
161	8	0.6	945	12	US-10-032-585-7114	Sequence 7114, App	234	7	0.5	108	10	US-09-764-864-817	Sequence 817, App

235	7	0.5	109	9	US-09-764-853-625	Sequence 625, App	308	7	0.5	229	15	US-10-325-878-9	Sequence 9, Appl
236	7	0.5	109	11	US-09-989-442-114	Sequence 114, App	309	7	0.5	231	14	US-10-124-557-30	Sequence 30, Appl
237	7	0.5	109	15	US-10-073-865-101	Sequence 101, App	310	7	0.5	238	15	US-10-195-117-3	Sequence 3, Appl
238	7	0.5	110	9	US-09-765-272-102	Sequence 102, App	311	7	0.5	238	15	US-10-195-117-4	Sequence 4, Appl
239	7	0.5	111	9	US-09-858-664-8	Sequence 8, Appl	312	7	0.5	242	12	US-10-296-658-17	Sequence 17, Appl
240	7	0.5	112	10	US-09-764-864-812	Sequence 812, App	313	7	0.5	243	11	US-09-533-029-14	Sequence 34, Appl
241	7	0.5	112	14	US-10-027-559-8	Sequence 8, Appl	314	7	0.5	243	11	US-09-934-455-16	Sequence 416, App
242	7	0.5	118	15	US-10-106-698-4760	Sequence 4760, App	315	7	0.5	246	9	US-09-864-761-36100	Sequence 36100, A
243	7	0.5	119	9	US-09-740-288A-10	Sequence 10, Appl	316	7	0.5	248	15	US-10-103-313-442	Sequence 442, App
244	7	0.5	120	15	US-10-101-464A-641	Sequence 641, App	317	7	0.5	256	15	US-10-156-761-11972	Sequence 11972, A
245	7	0.5	132	14	US-10-071-751-39	Sequence 39, Appl	318	7	0.5	259	11	US-09-934-455-298	Sequence 298, App
246	7	0.5	132	14	US-10-078-923-36	Sequence 36, Appl	319	7	0.5	260	11	US-09-906-333A-18	Sequence 18, App
247	7	0.5	134	10	US-09-738-626-5042	Sequence 5042, App	320	7	0.5	261	12	US-10-032-201B-162	Sequence 162, App
248	7	0.5	139	11	US-09-934-455-6	Sequence 6, Appl	321	7	0.5	261	14	US-10-062-254-140	Sequence 140, App
249	7	0.5	148	15	US-10-012-542-453	Sequence 453, App	322	7	0.5	261	14	US-10-062-254-142	Sequence 142, App
250	7	0.5	150	11	US-09-934-455-86	Sequence 86, Appl	323	7	0.5	262	11	US-09-906-333A-14	Sequence 14, Appl
251	7	0.5	152	11	US-09-974-879-542	Sequence 542, App	324	7	0.5	262	11	US-09-906-333A-16	Sequence 16, Appl
252	7	0.5	152	11	US-09-305-736-544	Sequence 544, App	325	7	0.5	270	14	US-10-078-929-38	Sequence 38, Appl
253	7	0.5	156	9	US-09-925-301-1154	Sequence 1154, App	326	7	0.5	270	15	US-10-286-254-56	Sequence 96, Appl
254	7	0.5	165	10	US-09-764-864-1267	Sequence 1267, App	327	7	0.5	273	9	US-09-815-242-5550	Sequence 5550, App
255	7	0.5	166	12	US-10-076-747-83	Sequence 83, App	328	7	0.5	282	15	US-10-295-403-54	Sequence 54, Appl
256	7	0.5	169	9	US-09-864-761-42972	Sequence 42972, A	329	7	0.5	284	11	US-10-056-019-3	Sequence 54, Appl
257	7	0.5	172	14	US-10-124-557-88	Sequence 88, Appl	330	7	0.5	284	11	US-09-934-455-104	Sequence 104, App
258	7	0.5	173	10	US-09-738-626-6015	Sequence 6015, App	331	7	0.5	285	15	US-10-157-031-132	Sequence 132, App
259	7	0.5	176	15	US-10-106-698-5506	Sequence 5906, App	332	7	0.5	286	9	US-09-815-242-12234	Sequence 12234, A
260	7	0.5	180	9	US-09-910-150-29	Sequence 29, Appl	333	7	0.5	288	12	US-09-896-186B-24	Sequence 24, Appl
261	7	0.5	181	14	US-10-062-254-369	Sequence 369, App	334	7	0.5	290	15	US-10-106-698-5388	Sequence 5388, App
262	7	0.5	181	15	US-10-149-819-14	Sequence 14, Appl	335	7	0.5	295	16	US-10-278-536-184	Sequence 184, App
263	7	0.5	185	12	US-10-342-224-80	Sequence 80, Appl	336	7	0.5	296	14	US-10-124-557-70	Sequence 70, Appl
264	7	0.5	187	12	US-09-903-190-102	Sequence 102, App	337	7	0.5	297	16	US-10-278-536-74	Sequence 74, Appl
265	7	0.5	191	11	US-09-934-455-54	Sequence 54, Appl	338	7	0.5	302	11	US-09-725-311-2	Sequence 2, Appl
266	7	0.5	192	10	US-09-764-868-1192	Sequence 1192, App	339	7	0.5	306	12	US-10-166-225A-101	Sequence 101, App
267	7	0.5	194	14	US-10-124-557-90	Sequence 90, Appl	340	7	0.5	307	12	US-10-238-075-453	Sequence 453, App
268	7	0.5	194	9	US-09-910-150-28	Sequence 28, Appl	341	7	0.5	307	15	US-10-278-173-148	Sequence 148, App
269	7	0.5	195	10	US-09-764-864-854	Sequence 854, App	342	7	0.5	307	16	US-10-278-536-52	Sequence 92, Appl
270	7	0.5	199	10	US-09-764-864-1291	Sequence 1291, App	343	7	0.5	310	15	US-10-007-521-22	Sequence 22, Appl
271	7	0.5	200	9	US-09-867-550-1934	Sequence 1934, App	344	7	0.5	311	12	US-10-259-165-318	Sequence 318, App
272	7	0.5	201	10	US-09-764-864-836	Sequence 836, App	345	7	0.5	312	10	US-09-738-626-5731	Sequence 5731, App
273	7	0.5	204	11	US-09-934-455-84	Sequence 84, Appl	346	7	0.5	313	12	US-09-896-186B-2	Sequence 2, Appl
274	7	0.5	204	14	US-10-124-557-92	Sequence 92, Appl	347	7	0.5	314	9	US-09-768-82-43	Sequence 43, Appl
275	7	0.5	207	10	US-09-965-602-34	Sequence 34, Appl	348	7	0.5	315	15	US-10-232-563-11	Sequence 11, Appl
276	7	0.5	207	11	US-09-975-719-211	Sequence 211, App	349	7	0.5	319	12	US-10-168-080-2	Sequence 2, Appl
277	7	0.5	207	14	US-10-062-254-174	Sequence 174, App	350	7	0.5	321	11	US-09-934-455-118	Sequence 418, App
278	7	0.5	208	14	US-09-733-507-13	Sequence 13, Appl	351	7	0.5	323	12	US-10-017-161-1620	Sequence 1620, App
279	7	0.5	208	14	US-10-124-557-132	Sequence 132, App	352	7	0.5	324	9	US-09-734-411-2	Sequence 2, Appl
280	7	0.5	209	10	US-10-124-557-94	Sequence 94, Appl	353	7	0.5	325	11	US-09-934-455-138	Sequence 138, App
281	7	0.5	211	10	US-09-908-805B-10	Sequence 10, Appl	354	7	0.5	328	10	US-09-900-557-2	Sequence 2, Appl
282	7	0.5	211	12	US-10-331-695-2	Sequence 2, Appl	355	7	0.5	328	15	US-10-157-031-12	Sequence 12, Appl
283	7	0.5	212	9	US-09-350-874-63	Sequence 63, Appl	356	7	0.5	330	11	US-09-725-311-4	Sequence 4, Appl
284	7	0.5	212	10	US-09-908-805B-4	Sequence 4, Appl	357	7	0.5	333	11	US-09-967-869A-14	Sequence 14, Appl
285	7	0.5	212	10	US-09-908-805B-12	Sequence 12, Appl	358	7	0.5	339	11	US-10-212-677-247	Sequence 247, App
286	7	0.5	212	12	US-10-331-695-4	Sequence 4, Appl	359	7	0.5	343	16	US-10-212-677-248	Sequence 248, App
287	7	0.5	212	15	US-10-038-010-60	Sequence 60, Appl	360	7	0.5	343	16	US-10-212-677-247	Sequence 247, App
288	7	0.5	212	15	US-10-278-173-122	Sequence 122, App	361	7	0.5	347	10	US-10-212-677-249	Sequence 249, App
289	7	0.5	212	15	US-10-295-403-16	Sequence 16, Appl	362	7	0.5	349	9	US-09-816-127-3	Sequence 3, Appl
290	7	0.5	212	15	US-10-106-698-63	Sequence 63, Appl	363	7	0.5	349	9	US-09-802-213-7	Sequence 7, Appl
291	7	0.5	214	9	US-09-925-302-649	Sequence 649, App	364	7	0.5	349	11	US-09-820-598-3	Sequence 3, Appl
292	7	0.5	214	12	US-10-293-971-16	Sequence 16, Appl	365	7	0.5	351	14	US-10-095-392-3	Sequence 3, Appl
293	7	0.5	216	12	US-10-267-430-2	Sequence 2, Appl	366	7	0.5	351	9	US-09-350-874-61	Sequence 61, Appl
294	7	0.5	217	14	US-10-124-557-76	Sequence 76, Appl	367	7	0.5	352	12	US-10-106-989-61	Sequence 61, Appl
295	7	0.5	217	15	US-10-021-811-63	Sequence 63, Appl	368	7	0.5	360	9	US-10-259-165-216	Sequence 216, App
296	7	0.5	217	15	US-10-106-698-5191	Sequence 5191, App	369	7	0.5	360	10	US-09-877-665-7	Sequence 7, Appl
297	7	0.5	217	15	US-10-106-698-5795	Sequence 5795, App	370	7	0.5	360	10	US-09-877-665-7	Sequence 7, Appl
298	7	0.5	218	11	US-09-934-455-192	Sequence 192, App	371	7	0.5	360	14	US-10-136-573A-7	Sequence 7, Appl
299	7	0.5	218	11	US-09-934-455-192	Sequence 192, App	371	7	0.5	360	15	US-10-215-862-7	Sequence 7, Appl
300	7	0.5	218	13	US-10-010-901-31	Sequence 31, Appl	372	7	0.5	368	9	US-09-805-550-4	Sequence 4, Appl
301	7	0.5	220	15	US-10-124-557-96	Sequence 96, Appl	373	7	0.5	368	9	US-09-801-368-308	Sequence 308, App
302	7	0.5	220	15	US-10-101-464A-628	Sequence 628, App	374	7	0.5	369	15	US-10-036-552-84	Sequence 84, Appl
303	7	0.5	221	12	US-10-032-585-7567	Sequence 7567, App	375	7	0.5	372	14	US-10-124-557-84	Sequence 64, Appl
304	7	0.5	222	15	US-10-156-761-7686	Sequence 7686, App	376	7	0.5	373	11	US-09-819-142-2	Sequence 2, Appl
305	7	0.5	222	15	US-10-156-761-12311	Sequence 12311, A	377	7	0.5	373	11	US-09-934-455-108	Sequence 108, App
306	7	0.5	226	11	US-09-967-869A-15	Sequence 15, Appl	378	7	0.5	379	9	US-09-740-288A-26	Sequence 26, Appl
307	7	0.5	229	12	US-10-411-224-150	Sequence 130, App	379	7	0.5	383	12	US-10-032-201B-196	Sequence 196, App
	7	0.5	229	15	US-10-091-333-9	Sequence 9, Appl	380	7	0.5	383	15	US-10-141-531-72	Sequence 72, Appl

381	7	0.5	383	15	US-10-141-531-87	Sequence 87, Appl	454	7	0.5	490	15	US-10-121-746-6	Sequence 6, Appl1
382	7	0.5	383	15	US-10-141-531-96	Sequence 96, Appl	455	7	0.5	491	11	US-09-999-220B-3	Sequence 3, Appl1
383	7	0.5	383	15	US-10-141-531-105	Sequence 105, Appl	456	7	0.5	491	11	US-09-999-220B-4	Sequence 4, Appl1
384	7	0.5	383	15	US-10-141-531-114	Sequence 114, Appl	457	7	0.5	491	12	US-10-032-588-7161	Sequence 7161, Ap
385	7	0.5	383	15	US-10-141-531-123	Sequence 123, Appl	458	7	0.5	495	12	US-10-293-086-73	Sequence 73, Appl
386	7	0.5	383	15	US-10-141-531-132	Sequence 132, Appl	459	7	0.5	500	15	US-10-006-852-16	Sequence 16, Appl
387	7	0.5	383	15	US-10-141-531-141	Sequence 141, Appl	460	7	0.5	500	15	US-10-219-220-149	Sequence 149, Appl
388	7	0.5	383	15	US-10-141-531-150	Sequence 150, Appl	461	7	0.5	500	15	US-10-228-897-13	Sequence 13, Appl
389	7	0.5	383	15	US-10-141-531-159	Sequence 159, Appl	462	7	0.5	503	10	US-09-801-368-424	Sequence 424, Appl
390	7	0.5	383	15	US-10-141-531-168	Sequence 168, Appl	463	7	0.5	506	10	US-09-976-165-19	Sequence 19, Appl
391	7	0.5	383	15	US-10-141-531-177	Sequence 177, Appl	464	7	0.5	506	12	US-10-342-276-19	Sequence 19, Appl
392	7	0.5	383	15	US-10-141-531-186	Sequence 186, Appl	465	7	0.5	511	11	US-09-928-523B-13	Sequence 13, Appl
393	7	0.5	383	15	US-10-141-531-195	Sequence 195, Appl	466	7	0.5	511	15	US-10-128-714-3150	Sequence 3150, Ap
394	7	0.5	383	15	US-10-141-531-204	Sequence 204, Appl	467	7	0.5	511	15	US-10-128-714-8150	Sequence 8150, Ap
395	7	0.5	384	10	US-09-945-249-8	Sequence 8, Appl1	468	7	0.5	511	15	US-10-156-761-9625	Sequence 9625, Ap
396	7	0.5	385	12	US-10-183-708-96	Sequence 96, Appl1	469	7	0.5	513	11	US-09-298-522B-12	Sequence 12, Appl
397	7	0.5	387	12	US-10-082-272-3	Sequence 96, Appl1	470	7	0.5	518	9	US-09-925-299-862	Sequence 862, Appl
398	7	0.5	391	15	US-10-156-275-6	Sequence 6, Appl1	471	7	0.5	518	11	US-09-925-299-862	Sequence 862, Appl
399	7	0.5	392	10	US-09-168-474-9	Sequence 9, Appl1	472	7	0.5	518	11	US-09-815-379-4	Sequence 4, Appl1
400	7	0.5	392	12	US-10-054-399A-9	Sequence 9, Appl1	473	7	0.5	518	15	US-10-283-023-2	Sequence 2, Appl1
401	7	0.5	392	15	US-10-060-036-67	Sequence 67, Appl	474	7	0.5	519	16	US-10-278-536-182	Sequence 182, Appl
402	7	0.5	395	11	US-09-949-427-2	Sequence 2, Appl1	475	7	0.5	522	15	US-10-081-051-107	Sequence 107, Appl
403	7	0.5	397	14	US-10-029-180-40	Sequence 40, Appl1	476	7	0.5	524	15	US-10-253-007-56	Sequence 56, Appl
404	7	0.5	398	10	US-09-961-527A-9	Sequence 9, Appl1	477	7	0.5	530	12	US-10-032-585-7699	Sequence 7699, Ap
405	7	0.5	399	11	US-09-934-455-278	Sequence 278, Appl	478	7	0.5	532	9	US-09-925-301-1004	Sequence 1004, Ap
406	7	0.5	401	12	US-09-949-029-146	Sequence 146, Appl	479	7	0.5	543	10	US-09-736-457-337	Sequence 337, Appl
407	7	0.5	403	9	US-09-802-213-5	Sequence 5, Appl1	480	7	0.5	543	10	US-09-902-941-337	Sequence 337, Appl
408	7	0.5	406	11	US-09-056-019-1	Sequence 1, Appl1	481	7	0.5	543	10	US-09-849-626-337	Sequence 337, Appl
409	7	0.5	411	15	US-10-156-761-11579	Sequence 1, Appl1	482	7	0.5	543	11	US-09-476-630-337	Sequence 337, Appl
410	7	0.5	414	9	US-09-802-213-3	Sequence 3, Appl1	483	7	0.5	543	12	US-10-251-661-6	Sequence 6, Appl1
411	7	0.5	415	10	US-09-931-457A-64	Sequence 64, Appl	484	7	0.5	543	12	US-10-113-872-337	Sequence 337, Appl
412	7	0.5	422	14	US-10-124-557-68	Sequence 68, Appl	485	7	0.5	543	12	US-10-032-585-7577	Sequence 7577, Ap
413	7	0.5	423	11	US-10-124-557-66	Sequence 66, Appl	486	7	0.5	543	15	US-10-017-754-337	Sequence 337, Appl
414	7	0.5	428	11	US-09-056-019-24	Sequence 24, Appl	487	7	0.5	543	15	US-10-205-822-95	Sequence 95, Appl
415	7	0.5	431	9	US-09-815-242-5406	Sequence 5406, Ap	488	7	0.5	544	12	US-10-032-585-7924	Sequence 7924, Ap
416	7	0.5	431	9	US-09-815-242-12488	Sequence 12488, A	489	7	0.5	553	12	US-10-347-272-13	Sequence 13, Appl
417	7	0.5	433	10	US-09-801-368-38	Sequence 38, Appl	490	7	0.5	553	12	US-10-106-699-5847	Sequence 5847, Ap
418	7	0.5	435	10	US-09-866-626-6703	Sequence 38, Appl	491	7	0.5	555	15	US-10-032-585-7914	Sequence 7914, Ap
419	7	0.5	438	10	US-09-738-626-6703	Sequence 6703, Ap	492	7	0.5	559	12	US-10-103-585-7914	Sequence 7914, Ap
420	7	0.5	446	12	US-10-254-995-6	Sequence 6, Appl1	493	7	0.5	562	15	US-10-083-357-1739	Sequence 1329, Ap
421	7	0.5	446	12	US-10-254-995-9	Sequence 9, Appl1	494	7	0.5	562	15	US-10-083-357-1739	Sequence 1329, Ap
422	7	0.5	447	12	US-10-082-272-2	Sequence 2, Appl1	495	7	0.5	578	12	US-10-032-585-7793	Sequence 7793, Ap
423	7	0.5	447	12	US-10-082-272-2	Sequence 2, Appl1	496	7	0.5	579	14	US-10-384-743-4	Sequence 4, Appl1
424	7	0.5	448	11	US-09-934-455-126	Sequence 126, Appl	497	7	0.5	579	14	US-10-158-895-4	Sequence 4, Appl1
425	7	0.5	450	9	US-09-826-752-2	Sequence 2, Appl1	498	7	0.5	582	15	US-10-122-805-3	Sequence 3, Appl1
426	7	0.5	450	15	US-10-153-668-268	Sequence 268, Appl	499	7	0.5	582	15	US-10-158-895-15	Sequence 15, Appl
427	7	0.5	450	15	US-10-153-668-380	Sequence 380, Appl	500	7	0.5	590	14	US-10-158-895-15	Sequence 15, Appl
428	7	0.5	450	15	US-10-153-668-380	Sequence 382, Appl	501	7	0.5	594	12	US-10-137-870-10	Sequence 10, Appl
429	7	0.5	452	9	US-09-815-242-10493	Sequence 10493, A	502	7	0.5	594	12	US-10-140-018-10	Sequence 10, Appl
430	7	0.5	455	12	US-09-848-841-12	Sequence 12, Appl	503	7	0.5	594	12	US-10-140-021-10	Sequence 10, Appl
431	7	0.5	459	11	US-09-056-019-39	Sequence 12, Appl	504	7	0.5	594	12	US-10-140-274-10	Sequence 10, Appl
432	7	0.5	460	10	US-09-925-300-1228	Sequence 1228, Ap	505	7	0.5	594	12	US-10-140-471-10	Sequence 10, Appl
433	7	0.5	460	11	US-09-056-019-38	Sequence 38, Appl	506	7	0.5	594	12	US-10-140-807-10	Sequence 10, Appl
434	7	0.5	463	14	US-10-124-557-54	Sequence 54, Appl	507	7	0.5	594	12	US-10-140-922-10	Sequence 10, Appl
435	7	0.5	464	11	US-09-770-509-22	Sequence 22, Appl	508	7	0.5	594	12	US-10-140-924-10	Sequence 10, Appl
436	7	0.5	465	9	US-09-884-260A-52	Sequence 52, Appl	509	7	0.5	594	12	US-10-140-926-10	Sequence 10, Appl
437	7	0.5	466	12	US-10-032-585-7466	Sequence 7466, Ap	510	7	0.5	594	12	US-10-141-698-10	Sequence 10, Appl
438	7	0.5	471	8	US-08-954-771-34	Sequence 34, Appl	511	7	0.5	594	12	US-10-141-702-10	Sequence 10, Appl
439	7	0.5	471	9	US-09-741-669-428	Sequence 428, Appl	512	7	0.5	594	12	US-10-141-704-10	Sequence 10, Appl
440	7	0.5	474	10	US-09-736-457-1812	Sequence 1812, Ap	513	7	0.5	594	12	US-10-142-421-10	Sequence 10, Appl
441	7	0.5	474	10	US-09-902-941-1812	Sequence 1812, Ap	514	7	0.5	594	12	US-10-142-432-10	Sequence 10, Appl
442	7	0.5	474	10	US-09-849-626-1812	Sequence 1812, Ap	515	7	0.5	594	12	US-10-142-432-10	Sequence 10, Appl
443	7	0.5	474	12	US-10-113-872-1812	Sequence 1812, Ap	516	7	0.5	594	12	US-10-143-033-10	Sequence 10, Appl
444	7	0.5	474	12	US-10-113-872-1812	Sequence 1812, Ap	517	7	0.5	594	12	US-10-143-033-10	Sequence 10, Appl
445	7	0.5	474	12	US-10-113-872-1812	Sequence 1812, Ap	518	7	0.5	594	12	US-10-144-994-10	Sequence 10, Appl
446	7	0.5	475	9	US-09-826-752-14	Sequence 381, Appl	519	7	0.5	594	12	US-10-145-628-10	Sequence 10, Appl
447	7	0.5	475	15	US-10-318-780-4	Sequence 14, Appl	520	7	0.5	594	12	US-10-145-633-10	Sequence 10, Appl
448	7	0.5	476	10	US-09-774-414-3	Sequence 3, Appl1	521	7	0.5	594	12	US-10-145-746-10	Sequence 10, Appl
449	7	0.5	477	12	US-10-161-051-43	Sequence 43, Appl1	522	7	0.5	594	12	US-10-145-823-10	Sequence 10, Appl
450	7	0.5	484	12	US-10-347-278-17	Sequence 17, Appl	523	7	0.5	594	12	US-10-145-826-10	Sequence 10, Appl
451	7	0.5	484	12	US-10-347-278-17	Sequence 17, Appl	524	7	0.5	594	12	US-10-145-826-10	Sequence 10, Appl
452	7	0.5	489	9	US-09-956-425-6	Sequence 6, Appl1	525	7	0.5	594	12	US-10-145-876-10	Sequence 10, Appl
453	7	0.5	489	11	US-09-533-029-60	Sequence 60, Appl	526	7	0.5	594	12	US-10-145-959-10	Sequence 10, Appl

673	7	0.5	594 15	US-10-121-045-10	Sequence 10, Appl	746	7	0.5	594 15	US-10-123-235-10	Sequence 10, Appl
674	7	0.5	594 15	US-10-123-292-10	Sequence 10, Appl	747	7	0.5	594 15	US-10-124-818-10	Sequence 10, Appl
675	7	0.5	594 15	US-10-123-903-10	Sequence 10, Appl	748	7	0.5	594 15	US-10-137-868-10	Sequence 10, Appl
676	7	0.5	594 15	US-10-124-819-10	Sequence 10, Appl	749	7	0.5	594 15	US-10-147-497-10	Sequence 10, Appl
677	7	0.5	594 15	US-10-124-822-10	Sequence 10, Appl	750	7	0.5	594 15	US-10-158-782-10	Sequence 10, Appl
678	7	0.5	594 15	US-10-140-925-10	Sequence 10, Appl	751	7	0.5	594 15	US-10-123-905-10	Sequence 10, Appl
679	7	0.5	594 15	US-10-160-498-10	Sequence 10, Appl	752	7	0.5	594 15	US-10-123-907-10	Sequence 10, Appl
680	7	0.5	594 15	US-10-124-824-10	Sequence 10, Appl	753	7	0.5	594 15	US-10-124-815-10	Sequence 10, Appl
681	7	0.5	594 15	US-10-127-825A-10	Sequence 10, Appl	754	7	0.5	594 15	US-10-125-928A-10	Sequence 10, Appl
682	7	0.5	594 15	US-10-127-829A-10	Sequence 10, Appl	755	7	0.5	594 15	US-10-127-828A-10	Sequence 10, Appl
683	7	0.5	594 15	US-10-127-835A-10	Sequence 10, Appl	756	7	0.5	594 15	US-10-127-821A-10	Sequence 10, Appl
684	7	0.5	594 15	US-10-127-839A-10	Sequence 10, Appl	757	7	0.5	594 15	US-10-127-822A-10	Sequence 10, Appl
685	7	0.5	594 15	US-10-127-901A-10	Sequence 10, Appl	758	7	0.5	594 15	US-10-127-832A-10	Sequence 10, Appl
686	7	0.5	594 15	US-10-128-693A-10	Sequence 10, Appl	759	7	0.5	594 15	US-10-127-826A-10	Sequence 10, Appl
687	7	0.5	594 15	US-10-131-818A-10	Sequence 10, Appl	760	7	0.5	594 15	US-10-127-827A-10	Sequence 10, Appl
688	7	0.5	594 15	US-10-131-818A-10	Sequence 10, Appl	761	7	0.5	594 15	US-10-127-828A-10	Sequence 10, Appl
689	7	0.5	594 15	US-10-131-823A-10	Sequence 10, Appl	762	7	0.5	594 15	US-10-127-830A-10	Sequence 10, Appl
690	7	0.5	594 15	US-10-131-824A-10	Sequence 10, Appl	763	7	0.5	594 15	US-10-127-832A-10	Sequence 10, Appl
691	7	0.5	594 15	US-10-131-830A-10	Sequence 10, Appl	764	7	0.5	594 15	US-10-127-833A-10	Sequence 10, Appl
692	7	0.5	594 15	US-10-131-837A-10	Sequence 10, Appl	765	7	0.5	594 15	US-10-127-834A-10	Sequence 10, Appl
693	7	0.5	594 15	US-10-137-872A-10	Sequence 10, Appl	766	7	0.5	594 15	US-10-127-836A-10	Sequence 10, Appl
694	7	0.5	594 15	US-10-147-500-10	Sequence 10, Appl	767	7	0.5	594 15	US-10-127-841A-10	Sequence 10, Appl
695	7	0.5	594 15	US-10-147-502-10	Sequence 10, Appl	768	7	0.5	594 15	US-10-127-844A-10	Sequence 10, Appl
696	7	0.5	594 15	US-10-147-515-10	Sequence 10, Appl	769	7	0.5	594 15	US-10-128-688A-10	Sequence 10, Appl
697	7	0.5	594 15	US-10-147-517-10	Sequence 10, Appl	770	7	0.5	594 15	US-10-128-689A-10	Sequence 10, Appl
698	7	0.5	594 15	US-10-147-526-10	Sequence 10, Appl	771	7	0.5	594 15	US-10-128-692A-10	Sequence 10, Appl
699	7	0.5	594 15	US-10-147-527-10	Sequence 10, Appl	772	7	0.5	594 15	US-10-128-694A-10	Sequence 10, Appl
700	7	0.5	594 15	US-10-121-041-10	Sequence 10, Appl	773	7	0.5	594 15	US-10-131-825A-10	Sequence 10, Appl
701	7	0.5	594 15	US-10-121-043-10	Sequence 10, Appl	774	7	0.5	594 15	US-10-230-417-10	Sequence 10, Appl
702	7	0.5	594 15	US-10-121-047-10	Sequence 10, Appl	775	7	0.5	594 15	US-10-131-815A-10	Sequence 10, Appl
703	7	0.5	594 15	US-10-123-215-10	Sequence 10, Appl	776	7	0.5	594 15	US-10-131-817A-10	Sequence 10, Appl
704	7	0.5	594 15	US-10-123-902-10	Sequence 10, Appl	777	7	0.5	594 15	US-10-131-821A-10	Sequence 10, Appl
705	7	0.5	594 15	US-10-123-908-10	Sequence 10, Appl	778	7	0.5	594 15	US-10-131-822A-10	Sequence 10, Appl
706	7	0.5	594 15	US-10-123-909-10	Sequence 10, Appl	779	7	0.5	594 15	US-10-131-828A-10	Sequence 10, Appl
707	7	0.5	594 15	US-10-123-910-10	Sequence 10, Appl	780	7	0.5	594 15	US-10-131-835A-10	Sequence 10, Appl
708	7	0.5	594 15	US-10-124-813-10	Sequence 10, Appl	781	7	0.5	594 15	US-10-137-864A-10	Sequence 10, Appl
709	7	0.5	594 15	US-10-124-817-10	Sequence 10, Appl	782	7	0.5	594 15	US-10-137-869A-10	Sequence 10, Appl
710	7	0.5	594 15	US-10-125-922-10	Sequence 10, Appl	783	7	0.5	594 15	US-10-147-523-10	Sequence 10, Appl
711	7	0.5	594 15	US-10-125-924-10	Sequence 10, Appl	784	7	0.5	594 15	US-10-158-785-10	Sequence 10, Appl
712	7	0.5	594 15	US-10-140-860-10	Sequence 10, Appl	785	7	0.5	594 15	US-10-121-051-10	Sequence 10, Appl
713	7	0.5	594 15	US-10-142-417-10	Sequence 10, Appl	786	7	0.5	594 15	US-10-121-042-10	Sequence 10, Appl
714	7	0.5	594 15	US-10-147-519-10	Sequence 10, Appl	787	7	0.5	594 15	US-10-123-912-10	Sequence 10, Appl
715	7	0.5	594 15	US-10-157-782-10	Sequence 10, Appl	788	7	0.5	594 15	US-10-192-007-10	Sequence 10, Appl
716	7	0.5	594 15	US-10-152-395-10	Sequence 10, Appl	789	7	0.5	594 15	US-10-194-359-10	Sequence 10, Appl
717	7	0.5	594 15	US-10-125-926A-10	Sequence 10, Appl	790	7	0.5	594 15	US-10-127-847A-10	Sequence 10, Appl
718	7	0.5	594 15	US-10-125-930A-10	Sequence 10, Appl	791	7	0.5	594 15	US-10-175-590-10	Sequence 10, Appl
719	7	0.5	594 15	US-10-127-831A-10	Sequence 10, Appl	792	7	0.5	594 16	US-10-137-866-10	Sequence 10, Appl
720	7	0.5	594 15	US-10-127-837A-10	Sequence 10, Appl	793	7	0.5	594 16	US-10-146-726-10	Sequence 10, Appl
721	7	0.5	594 15	US-10-127-838B-10	Sequence 10, Appl	794	7	0.5	594 16	US-10-146-727-10	Sequence 10, Appl
722	7	0.5	594 15	US-10-127-842A-10	Sequence 10, Appl	795	7	0.5	594 16	US-10-146-788-10	Sequence 10, Appl
723	7	0.5	594 15	US-10-127-843A-10	Sequence 10, Appl	796	7	0.5	594 16	US-10-152-380-10	Sequence 10, Appl
724	7	0.5	594 15	US-10-127-845A-10	Sequence 10, Appl	797	7	0.5	594 16	US-10-153-934-10	Sequence 10, Appl
725	7	0.5	594 15	US-10-127-846A-10	Sequence 10, Appl	798	7	0.5	9	US-09-815-242-10990	Sequence 10990, A
726	7	0.5	594 15	US-10-127-848A-10	Sequence 10, Appl	799	7	0.5	613 12	US-09-848-841-16	Sequence 16, Appl
727	7	0.5	594 15	US-10-127-849A-10	Sequence 10, Appl	800	7	0.5	646 9	US-09-815-242-10311	Sequence 10311, A
728	7	0.5	594 15	US-10-127-850A-10	Sequence 10, Appl	801	7	0.5	646 9	US-09-815-242-13769	Sequence 13769, A
729	7	0.5	594 15	US-10-127-851A-10	Sequence 10, Appl	802	7	0.5	653 12	US-10-199-672-402	Sequence 402, App
730	7	0.5	594 15	US-10-128-684A-10	Sequence 10, Appl	803	7	0.5	653 12	US-10-187-749-402	Sequence 402, App
731	7	0.5	594 15	US-10-128-686A-10	Sequence 10, Appl	804	7	0.5	653 12	US-10-194-457-402	Sequence 402, App
732	7	0.5	594 15	US-10-128-690A-10	Sequence 10, Appl	805	7	0.5	653 12	US-10-184-642-402	Sequence 402, App
733	7	0.5	594 15	US-10-128-691A-10	Sequence 10, Appl	806	7	0.5	653 12	US-10-196-747-402	Sequence 402, App
734	7	0.5	594 15	US-10-131-819A-10	Sequence 10, Appl	807	7	0.5	653 12	US-10-173-689-402	Sequence 402, App
735	7	0.5	594 15	US-10-131-829A-10	Sequence 10, Appl	808	7	0.5	653 12	US-10-173-690-402	Sequence 402, App
736	7	0.5	594 15	US-10-131-836A-10	Sequence 10, Appl	809	7	0.5	653 12	US-10-173-691-402	Sequence 402, App
737	7	0.5	594 15	US-10-146-729-10	Sequence 10, Appl	810	7	0.5	653 12	US-10-173-692-402	Sequence 402, App
738	7	0.5	594 15	US-10-146-791-10	Sequence 10, Appl	811	7	0.5	653 12	US-10-173-694-402	Sequence 402, App
739	7	0.5	594 15	US-10-147-484-10	Sequence 10, Appl	812	7	0.5	653 12	US-10-173-698-402	Sequence 402, App
740	7	0.5	594 15	US-10-147-508-10	Sequence 10, Appl	813	7	0.5	653 12	US-10-173-699-402	Sequence 402, App
741	7	0.5	594 15	US-10-147-512-10	Sequence 10, Appl	814	7	0.5	653 12	US-10-173-707-402	Sequence 402, App
742	7	0.5	594 15	US-10-175-735-10	Sequence 10, Appl	815	7	0.5	653 12	US-10-174-569-402	Sequence 402, App
743	7	0.5	594 15	US-10-121-040-10	Sequence 10, Appl	816	7	0.5	653 12	US-10-174-583-402	Sequence 402, App
744	7	0.5	594 15	US-10-121-056-10	Sequence 10, Appl	817	7	0.5	653 12	US-10-174-587-402	Sequence 402, App
745	7	0.5	594 15	US-10-121-061-10	Sequence 10, Appl	818	7	0.5	653 12	US-10-174-589-402	Sequence 402, App

965	7	0.5	653	15	US-10-188-773-402	Sequence 402, App
966	7	0.5	653	15	US-10-188-781-402	Sequence 402, App
967	7	0.5	653	15	US-10-194-361-402	Sequence 402, App
968	7	0.5	653	15	US-10-194-423-402	Sequence 402, App
969	7	0.5	653	15	US-10-195-897-402	Sequence 402, App
970	7	0.5	653	15	US-10-195-901-402	Sequence 402, App
971	7	0.5	653	15	US-10-195-902-402	Sequence 402, App
972	7	0.5	653	15	US-10-196-743-402	Sequence 402, App
973	7	0.5	653	15	US-10-196-760-402	Sequence 402, App
974	7	0.5	653	15	US-10-173-708-402	Sequence 402, App
975	7	0.5	653	15	US-10-176-479-402	Sequence 402, App
976	7	0.5	653	15	US-10-176-748-402	Sequence 402, App
977	7	0.5	653	15	US-10-176-916-402	Sequence 402, App
978	7	0.5	653	15	US-10-179-507-402	Sequence 402, App
979	7	0.5	653	15	US-10-179-516-402	Sequence 402, App
980	7	0.5	653	15	US-10-179-519-402	Sequence 402, App
981	7	0.5	653	15	US-10-179-525-402	Sequence 402, App
982	7	0.5	653	15	US-10-180-540-402	Sequence 402, App
983	7	0.5	653	15	US-10-180-545-402	Sequence 402, App
984	7	0.5	653	15	US-10-183-006-402	Sequence 402, App
985	7	0.5	653	15	US-10-183-008-402	Sequence 402, App
986	7	0.5	653	15	US-10-183-017-402	Sequence 402, App
987	7	0.5	653	15	US-10-183-019-402	Sequence 402, App
988	7	0.5	653	15	US-10-184-618-402	Sequence 402, App
989	7	0.5	653	15	US-10-184-625-402	Sequence 402, App
990	7	0.5	653	15	US-10-184-626-402	Sequence 402, App
991	7	0.5	653	15	US-10-184-627-402	Sequence 402, App
992	7	0.5	653	15	US-10-184-645-402	Sequence 402, App
993	7	0.5	653	15	US-10-184-654-402	Sequence 402, App
994	7	0.5	653	15	US-10-184-655-402	Sequence 402, App
995	7	0.5	653	15	US-10-188-774-402	Sequence 402, App
996	7	0.5	653	15	US-10-188-775-402	Sequence 402, App
997	7	0.5	653	15	US-10-194-462-402	Sequence 402, App
998	7	0.5	653	15	US-10-196-745-402	Sequence 402, App
999	7	0.5	653	15	US-10-196-762-402	Sequence 402, App
1000	7	0.5	653	15	US-10-197-695-402	Sequence 402, App

ALIGNMENTS

RESULT 1
US-10-006-780-2
Sequence 2, Application US/10006780
Publication No. US20030104496A1
GENERAL INFORMATION:
APPLICANT: Sakowicz, Roman
APPLICANT: Beraud, Christophe
APPLICANT: Guo, Jun
APPLICANT: Freedman, Richard
TITLE OF INVENTION: NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND
TITLE OF INVENTION: METHODS FOR ITS USE
FILE REFERENCE: CYTOPO83
CURRENT APPLICATION NUMBER: US/10/006,780
CURRENT FILING DATE: 2001-11-30
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1288
TYPE: PRT
ORGANISM: Plasmodium falciparum
US-10-006-780-2

Query Match 100.0%; Score 1288; DB 15; Length 1288;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNSKIKVYVRKRPSELKKKKSDIITVKNKCTLYIDEPYKVMYKIRHHEFTYDKV 60
DB 1 MNSKIKVYVRKRPSELKKKKSDIITVKNKCTLYIDEPYKVMYKIRHHEFTYDKV 60
QY 61 FDDTVNFTYVENTIKPLIIDLYENGCVCSFAYGQTSGKTYTMLGSQPYGSDTPGIF 120
DB 61 FDDTVNFTYVENTIKPLIIDLYENGCVCSFAYGQTSGKTYTMLGSQPYGSDTPGIF 120

Db	61	FDDTVNFTYVENTIKPLIIDLYENGCVCSFAYGQTSGKTYTMLGSQPYGSDTPGIF	120
QY	121	QYAGDIFFETIYDKONTGIFISFEYICGALYDLOKRVAALENKEKEVYDKL	180
Db	121	QYAGDIFFETIYDKONTGIFISFEYICGALYDLOKRVAALENKEKEVYDKL	180
QY	181	ILVLTEBELILMDGVLLRKIGVSONDESSRAHILNIDKIDNKTSIACKIAEIDL	240
Db	181	ILVLTEBELILMDGVLLRKIGVSONDESSRAHILNIDKIDNKTSIACKIAEIDL	240
QY	241	AGSERGADYTSQNKQOTGDANINRSILAKCEIRAMDSPKNIIPRDSLTLYLRIEY	300
Db	241	AGSERGADYTSQNKQOTGDANINRSILAKCEIRAMDSPKNIIPRDSLTLYLRIEY	300
QY	301	GKSKSIIMANISPTISCEQTLNLTLRSSVFKPKFKSTINEEDPTERTISILDSKGS	360
Db	301	GKSKSIIMANISPTISCEQTLNLTLRSSVFKPKFKSTINEEDPTERTISILDSKGS	360
QY	361	EMASSIENVVIKSNHLSNNNNKINRGKINDKIERNNILKXKSPDKREGFTSTFGKY	420
Db	361	EMASSIENVVIKSNHLSNNNNKINRGKINDKIERNNILKXKSPDKREGFTSTFGKY	420
QY	421	SSLANDIKIKKKKKGLINYKSTLYNDNTINKGHNNNNNNNDNNNNNNNDSS	480
Db	421	SSLANDIKIKKKKKGLINYKSTLYNDNTINKGHNNNNNNNDNNNNNNNDSS	480
QY	481	SMVNMNMHMINNNINNVNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNDSS	540
Db	481	SMVNMNMHMINNNINNVNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNDSS	540
QY	541	HLNNDKSIPLAKKNLJDNITKLXNRSSCDNIMKKKNLILARHSVSKLTWFSYDQKN	600
Db	541	HLNNDKSIPLAKKNLJDNITKLXNRSSCDNIMKKKNLILARHSVSKLTWFSYDQKN	600
QY	601	KDNTFFSKINAKMEDNTPKDIILYESRVSVMNGVLLGLNKQTHHDISTDEHNDKIN	660
Db	601	KDNTFFSKINAKMEDNTPKDIILYESRVSVMNGVLLGLNKQTHHDISTDEHNDKIN	660
QY	661	NGVNIINNNSNNSINNSNNSINNSNNSINNSNNSINNSNNSINNSNNSINNSNNS	720
Db	661	NGVNIINNNSNNSINNSNNSINNSNNSINNSNNSINNSNNSINNSNNSINNSNNS	720
QY	721	NNDNIPFDALSCDNNMYPNITNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	780
Db	721	NNDNIPFDALSCDNNMYPNITNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	780
QY	781	HNLFPDNNKQTSNIQININNNKQDGNVYSNMFCYINLNDKYLIDLNNKQKQKNIH	840
Db	781	HNLFPDNNKQTSNIQININNNKQDGNVYSNMFCYINLNDKYLIDLNNKQKQKNIH	840
QY	841	GCDNNIIONNDPEKKKKTFFYNNNNNVIYNNNNGNNSPRMKYGLGSHTSIDMKNE	900
Db	841	GCDNNIIONNDPEKKKKTFFYNNNNNVIYNNNNGNNSPRMKYGLGSHTSIDMKNE	900
QY	901	MKNMEMDNEMKONHISKNNNSSSSSSSNNNIYNNINDDTFONDYCHANDFTIRKN	960
Db	901	MKNMEMDNEMKONHISKNNNSSSSSSSNNNIYNNINDDTFONDYCHANDFTIRKN	960
QY	961	NTNINSNIYONDDIITYINSINDYMSNTLLHFKEKYPTLSTNEDLYNKEMEKHLRLD	1020
Db	961	NTNINSNIYONDDIITYINSINDYMSNTLLHFKEKYPTLSTNEDLYNKEMEKHLRLD	1020
QY	1021	DODKYDDNNNVNKNKNVNDNNVNDNNVNDNNVNDNNVNDNNVNDNNVNDNNVNDNN	1080
Db	1021	DODKYDDNNNVNKNKNVNDNNVNDNNVNDNNVNDNNVNDNNVNDNNVNDNNVNDNN	1080
QY	1081	NFNNEVLSYFOKQVDTIINNCLNSLDISSYVDTEKILNNILSKYAEKQVYKXYIN	1140
Db	1081	NFNNEVLSYFOKQVDTIINNCLNSLDISSYVDTEKILNNILSKYAEKQVYKXYIN	1140
QY	1141	EDIKMSLEBIDTAQSIYERKRVLLTKLALLFKQAVDTQINNETSULRQDLVWCHICNN	1200
Db	1141	EDIKMSLEBIDTAQSIYERKRVLLTKLALLFKQAVDTQINNETSULRQDLVWCHICNN	1200

QY	1201	NPDDQFHYAARLEKDIINIMLRQIACEENIRLYQFLAVEYONKSASVLLANYSN	1260
Db	1201	NPDDQFHYAARLEKDIINIMLRQIACEENIRLYQFLAVEYONKSASVLLANYSN	1260
QY	1261	NGDIITLLANKKLVDNIRKSMDSMHNNIHKK	1288
Db	1261	NGDIITLLANKKLVDNIRKSMDSMHNNIHKK	1288

```

RESULT 2
US-10-006-780-6
; Sequence 6, Application US/10006780
; Publication No. US2003010496A1
; GENERAL INFORMATION:
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; APPLICANT: Guo, Jun
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND
; TITLE OF INVENTION: METHODS FOR ITS USE
; FILE REFERENCE: CYTO083
; CURRENT APPLICATION NUMBER: US/10/006,780
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 355
; TYPE: PR1
; ORGANISM: P. Falciparum
; US-10-006-780-6

```

Query Match	27.5%	Score 354	DB 15	Length 355
Best Local Similarity	100.0%	Pred. No. 0		
Matches 354	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	4	KIKVAVKRPRLSELEKKKDSDIITVKNKCTLYIDEPKYKDMTKYIERHBEIVDYKVD	63	
Db	2	KIKVAVKRPRLSELEKKKDSDIITVKNKCTLYIDEPKYKDMTKYIERHBEIVDYKVD	61	
QY	64	TVDNFTIYENTIKRLIIDLYENGCVCSPAYGQTGSGKTYTMLGSPYGGSDTPIGFOYA	123	
Db	62	TVDNFTIYENTIKRLIIDLYENGCVCSPAYGQTGSGKTYTMLGSPYGGSDTPIGFOYA	121	
QY	124	AGDIFTEFLNYDKONTKGIFISFEIYICGKLYDLQKRWAAALENKKVAVVVDKILR	183	
Db	122	AGDIFTEFLNYDKONTKGIFISFEIYICGKLYDLQKRWAAALENKKVAVVVDKILR	181	
QY	184	VLTEKEELIKNKIDVLRKIGVNSQNDSSRSHAILNIDDKDINKNTSLGKAIFIDLAGS	243	
Db	182	VLTEKEELIKNKIDVLRKIGVNSQNDSSRSHAILNIDDKDINKNTSLGKAIFIDLAGS	241	
QY	244	ERGADTVSQNKQOTDGDANINRSLLALKECIRAMDSDKNHPFRDSELTXYLRDIPYQKS	303	
Db	242	ERGADTVSQNKQOTDGDANINRSLLALKECIRAMDSDKNHPFRDSELTXYLRDIPYQKS	301	
QY	304	KSLMINAISPISCCBOTLNTLRSSRKYKPFKXKSTCINEDDPTNTERISILDS	357	
Db	302	KSLMINAISPISCCBOTLNTLRSSRKYKPFKXKSTCINEDDPTNTERISILDS	355	

RESULT 3
US-10-006-780-4
Sequence 4, Application US/10006780
Publication No. US2003010496A1
GENERAL INFORMATION:
APPLICANT: Sakowicz, Roman
APPLICANT: Berard, Christophe
APPLICANT: Guo, Jun
APPLICANT: Freedman, Richard
TITLE OF INVENTION: NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND
TITLE OF INVENTION: METHODS FOR ITS USE
FILE REFERENCE: CYTO083

```

; CURRENT APPLICATION NUMBER: US/10/006,780
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ. ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ. ID NO. 4
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; OS-10-006-780-4

```

Query Match	25.8%	Score 332	DB 15	Length 332
Best Local Similarity	100.0%	Pred. No. 2e-300		
Matches 332	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	4	KIKVVRKRPRLSELSEKKKKDSIIITVKNNTLYIDBPRYKVDMTKYIERHFFIVDKYFDD	63	
Db	1	KIKVVRKRPRLSELSEKKKKDSIIITVKNNTLYIDBPRYKVDMTKYIERHFFIVDKYFD	60	
QY	64	TVDNFTYYENTIKPLIIDLYENGCVCSCPAYGOTGSGKTYTMLGSDPYGOSDFGIFQYA	123	
Db	61	TVNFTYYENTIKPLIIDLYENGCVCSCPAYGOTGSGKTYTMLGSDPYGOSDFGIFQYA	120	
QY	124	AGDIFPEFLNIYDDKNKGI FIFSELEYCGALYDLOKRVVAALLENKKEKVEVVDLKLIR	183	
Db	121	AGDIFPEFLNIYDDKNKGI FIFSELEYCGALYDLOKRVVAALLENKKEKVEVVDLKLIR	180	
QY	184	VLTREELILMIDGVLRLKIGVNSONDESSRSHAILNI DLKDINKNTSLGKIAFIDLAGS	243	
Db	181	VLTREELILMIDGVLRLKIGVNSONDESSRSHAILNI DLKDINKNTSLGKIAFIDLAGS	240	
QY	244	ERGADTVYSQNKQFOTGDGANINRSLLALKECIRAMDSDKNHI PRDSELTIVYLRDIFVGS	303	
Db	241	ERGADTVYSQNKQFOTGDGANINRSLLALKECIRAMDSDKNHI PRDSELTIVYLRDIFVGS	300	
QY	304	KSIMIANISPTISCCBOTNLTLAYSAVRVKNK	335	
Db	301	KSIMIANISPTISCCBOTNLTLAYSSVRVKNK	332	

```

RESULT 4
US-10-006-780-10
; Sequence 10. Application US/10006780
; Publication No. US20030104496A1
; GENERAL INFORMATION:
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; APPLICANT: Guo, Jun
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND
; TITLE OF INVENTION: METHODS FOR ITS USE
; FILE REFERENCE: CITOPO83
; CURRENT APPLICATION NUMBER: US/10/006,780
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 332
; TYPE: PRT
; ORGANISM: P. Falciparum
; US-10-006-780-10

```

Query Match	25.8%	Score 332	DB 15	Length 333
Best Local Similarity	100.0%	Pred. No. 20-300		
Matches 333	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	5	IKVAVKRPRLSELKKKKSDITLVKONCLTIDEPBRYKVDMMKTYIERHFIYDKVPDDT	64	
Db	1	IKVAVKRPRLSELKKKKSDITLVKONCLTIDEPBRYKVDMMKTYIERHFIYDKVPDDT	60	
QY	65	VDNFTYVENTIKPLIIDLYENGCVCFAVGQTGSGKTYTLMGSGPYGSDTGGIFQYAA	124	
Db	61	VDNFTYVENTIKPLIIDLYENGCVCFAVGQTGSGKTYTLMGSGPYGSDTGGIFQYAA	120	

Query Match	25.6%;	Score 330;	DB 15;	Length 361;
Best Local Similarity	100.0%;	Pred. No. 1.6e-298;		
Matches 330;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	4	KIKVVRKRLPLSELEKKKKDSDIITVKNNTLYIDEPRYAVDMTKYIERHEFIVDKYVDD	63	
Db	2	KIKVVRKRLPLSELEKKKKDSDIITVKNNTLYIDEPRYAVDMTKYIERHEFIVDKYVDD	61	
QY	64	TYDNFTVVENTIRPLIIDLYENGCVCSCPAYGQGSQKTYTMLGSGQYTGSDTFGIRQYA	122	
Db	62	TYDNFTVVENTIRPLIIDLYENGCVCSCPAYGQGSQKTYTMLGSGQYTGSDTFGIRQYA	122	
QY	124	AGDIFFTFLNIYDKONTKGFISFEYEICGLYDILQSRKNVAALENKKEVAVVDLKITLR	183	
Db	122	AGDIFFTFLNIYDKONTKGFISFEYEICGLYDILQSRKNVAALENKKEVAVVDLKITLR	183	
QY	184	VLTKEELLKMLDGVLLRKIGVNSONDDESSRSHALINIDILDKINTKSTLGKIAFIDLAGS	243	
Db	182	VLTKEELLKMLDGVLLRKIGVNSONDDESSRSHALINIDILDKINTKSTLGKIAFIDLAGS	243	
QY	244	ERGAUTVSONKQOTOTDGANINRSLALKECIRAMDSJDNHI PPFDSLETIKYLRI FVFGKS	303	
Db	242	ERGAUTVSONKQOTOTDGANINRSLALKECIRAMDSJDNHI PPFDSLETIKYLRI FVFGKS	303	
QY	304	KSIMIANISPTISCCBOTLNTLTLYSSSVKN	333	
Db	302	KSIMIANISPTISCCBOTLNTLTLRSSSVKN	331	

```

APPLICANT: Busby, Robert
APPLICANT: Cali, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: No. US20020128250A1man, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amit
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801.368
PRIOR FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/160,587
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 36
LENGTH: 666
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-801-368-36

Query Match 1.3%; Score 17; DB 10; Length 666;
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNNNT 758
|||
Db 351 NNNNNNNNNNNNNNNNNNI 367

RESULT 7
US-10-032-585-7714
; Sequence 7714, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 7714
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7714

Query Match 1.2%; Score 16; DB 12; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 741 TNNNNNNNNNNNNNNNNN 756
|||||
Db 172 TNNNNNNNNNNNNNNNNN 187

RESULT 8
US-09-994-485-8
; Sequence 8, Application US/09994485
; Patent No. US20020142429A1
; GENERAL INFORMATION:

```


APPLICANT: Ryzanov, Alexey G.
Halt, William N.
Pavur, Karen S.
TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/994,485
FILING DATE: 27-No. US20020142429A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/914,999
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 601-1-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 732 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Dictyostelium discoideum
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-994-485-8
Query Match 1.2%; Score 16; DB 10; Length 732;
Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 NNNNNNNNNNNNNNNN 757
DB 355 NNNNNNNNNNNNNNNN 370
RESULT 9
US-09-832-292-12
Sequence 12, Application US/09832292
Patent No. US2002017205A1
GENERAL INFORMATION:
APPLICANT: Ryzanov, Alexey
TITLE OF INVENTION: MAMMALIAN ALPHA-KINASE PROTEINS, NUCLEIC ACIDS AND
TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USES THEREOF
FILE REFERENCE: 601-1-098CIP
CURRENT APPLICATION NUMBER: US/09/832,292
CURRENT FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 09/632,131
PRIOR FILING DATE: 2001-08-03
NUMBER OF SEQ ID NOS: 45
SOFTWARE: Patent In version 3.1
SEQ ID NO 12
LENGTH: 732
TYPE: PRT

ORGANISM: Dictyostelium discoideum
US-09-832-292-12
Query Match 1.2%; Score 16; DB 10; Length 732;
Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 NNNNNNNNNNNNNNNN 757
DB 355 NNNNNNNNNNNNNNNN 370
RESULT 10
US-09-801-368-224
Sequence 224, Application US/09801368
Patent No. US20020128250A1
GENERAL INFORMATION:
APPLICANT: Busby, Robert
APPLICANT: Cali, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: No. US20020128250A1man, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Method for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801,368
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/160,587
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: Patent In version 3.0
SEQ ID NO 224
LENGTH: 738
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-801-368-224
Query Match 1.2%; Score 16; DB 10; Length 758;
Best Local Similarity 100.0%; Pred. No. 5.4e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 NNNNNNNNNNNNNNNN 757
DB 605 NNNNNNNNNNNNNNNN 620
RESULT 11
US-10-032-585-7565
Sequence 7565, Application US/10032585
Patent No. US20030180953A1
GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jiang
APPLICANT: Charles, Boone
APPLICANT: Howard, Bussey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032,585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: Patent In version 3.1
SEQ ID NO 7565
LENGTH: 798
TYPE: PRT


```
; ORGANISM: Candida albicans
US-10-032-585-7565

Query Match
Best Local Similarity 100.0%; Score 16; DB 12; Length 798;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNN 757
DB 390 NNNNNNNNNNNNNNNNN 405

RESULT 12
US-09-801-368-132
; Sequence 132, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amit
; APPLICANT: Silva, Jelf
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132
; LENGTH: 961
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-132

Query Match
Best Local Similarity 100.0%; Score 16; DB 10; Length 961;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNN 757
DB 515 NNNNNNNNNNNNNNNNN 530

RESULT 13
US-10-087-464-50
; Sequence 50, Application US/10087464
; Publication No. US20030059436A1
; GENERAL INFORMATION:
; APPLICANT: Chishti, Alhar
; APPLICANT: Oh, Steven
; APPLICANT: Liu, David
; APPLICANT: Goel, Vikas
; APPLICANT: Li, Xuerong
; TITLE OF INVENTION: Band 3 Antigenic Peptides, Malaria Polypeptides and Uses Thereof
; FILE REFERENCE: S1237/7019
; CURRENT APPLICATION NUMBER: US/10/087,464
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 06/272,930
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 50
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-087-464-50

Query Match
Best Local Similarity 100.0%; Score 16; DB 15; Length 1331;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNN 757
DB 1178 NNNNNNNNNNNNNNNNN 1193

RESULT 14
US-10-135-322-17
; Sequence 17, Application US/10135322
; Publication No. US20020173017A1
; GENERAL INFORMATION:
; APPLICANT: BENEFY, PN
; APPLICANT: HELARUTTA, Y
; APPLICANT: MAHONEN, AP
; APPLICANT: BONKE, AMW
; APPLICANT: KAUPPINEN, L
; APPLICANT: RIKONEN, M
; TITLE OF INVENTION: WOODEN LEG GENE, PROMOTER AND USES THEREOF
; FILE REFERENCE: 5914-086-999
; CURRENT APPLICATION NUMBER: US/10/135,322
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: 60/253,739
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 2150
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-135-322-17

Query Match
Best Local Similarity 100.0%; Score 16; DB 14; Length 2150;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNN 757
DB 105 NNNNNNNNNNNNNNNNN 120

RESULT 15
US-10-304-095-6
; Sequence 6, Application US/10304095
; Publication No. US20030134275A1
; GENERAL INFORMATION:
; APPLICANT: Long, David M.
; APPLICANT: Metz, Anneke M.
; APPLICANT: Love, Ruschelle A.
; TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes
; FILE REFERENCE: 47714-5009-US
; CURRENT APPLICATION NUMBER: US/10/304,095
; PRIOR FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: US/09/417,485
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2184
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
FEATURES:
NAME/KEY: unsure
LOCATION: (330)..(335)
OTHER INFORMATION: Xaa at position 330 = Leu or Ile;
```



```
; OTHER INFORMATION: Xaa at position 335 = Asp or Gly.
US-10-304-095-6
Query Match 1.2%; Score 16; DB 12; Length 2184;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNN 757
DB 239 NNNNNNNNNNNNNNNNN 254

RESULT 16
US-10-184-644-35
; Sequence 35, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 35
; LENGTH: 2690
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 2039-2065
; OTHER INFORMATION: unknown base
US-10-184-644-35

Query Match 1.2%; Score 16; DB 15; Length 2690;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNN 757
DB 2039 NNNNNNNNNNNNNNNNN 2054

RESULT 17
US-10-184-634-35
; Sequence 35, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C217
```

```
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 35
; LENGTH: 2690
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 2039-2065
; OTHER INFORMATION: unknown base
US-10-184-634-35

Query Match 1.2%; Score 16; DB 15; Length 2690;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNN 757
DB 2039 NNNNNNNNNNNNNNNNN 2054

RESULT 18
US-09-864-761-47535
; Sequence 47535, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
```



```

; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 47535
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010436.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 46
; OTHER INFORMATION: EST_HUMAN HIT: BF327361.1, EVALU 2.20e+00
US-09-864-761-47535

Query Match          1.2%; Score 15; DB 9; Length 73;
Best Local Similarity 100.0%; Pred. No. 5.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      742 NNNNNNNNNNNNNNNN 756
Db      11 NNNNNNNNNNNNNNNN 25

RESULT 19
US-10-422-866-4
; Sequence 4, Application US/10422866
; Publication No. US20030170857A1
; GENERAL INFORMATION:
; APPLICANT: Berand, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. US20030170857A1 motor proteins and methods for
; FILE REFERENCE: 1054
; CURRENT APPLICATION NUMBER: US/10/422,866
; CURRENT FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: US/09/722,129
; PRIOR FILING DATE: 2000-11-24
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Human
US-10-422-866-4

Query Match          1.2%; Score 15; DB 12; Length 303;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      91 CFAYGQTSGKTYTM 105
Db      90 CFAYGQTSGKTYTM 104

RESULT 20
US-10-422-866-2
; Sequence 2, Application US/10422866
; Publication No. US20030170857A1
; GENERAL INFORMATION:
; APPLICANT: Berand, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. US20030170857A1 motor proteins and methods for
; FILE REFERENCE: 1054
; CURRENT APPLICATION NUMBER: US/10/422,866
; CURRENT FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: US/09/722,129
; PRIOR FILING DATE: 2000-11-24
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Human
US-10-422-866-2
```

```

Query Match          1.2%; Score 15; DB 12; Length 492;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      91 CFAYGQTSGKTYTM 105
Db      279 CFAYGQTSGKTYTM 293

RESULT 21
US-10-304-095-8
; Sequence 8, Application US/10304095
; Publication No. US20030134275A1
; GENERAL INFORMATION:
; APPLICANT: Long, David M.
; APPLICANT: Metz, Anneke M.
; APPLICANT: Love, Ruschelle A.
; TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes
; FILE REFERENCE: 47714-5009-US
; CURRENT APPLICATION NUMBER: US/10/304,095
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: US/09/417,485
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 794
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-304-095-8

Query Match          1.2%; Score 15; DB 12; Length 794;
Best Local Similarity 100.0%; Pred. No. 4.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      744 NNNNNNNNNNNNNNNN 758
Db      130 NNNNNNNNNNNNNNNN 144

RESULT 22
US-09-826-752-6
; Sequence 6, Application US/09826752
; Patent No. US20010026930A1
; GENERAL INFORMATION:
; APPLICANT: Guarente, Leonard P.
; APPLICANT: Austriaco Jr., Nicanor
; APPLICANT: Claus, James J.
; APPLICANT: Cole, Francesca
; APPLICANT: Kennedy, Brian
; TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESCENCE IN
; FILE REFERENCE: 0050.1491-005
; CURRENT APPLICATION NUMBER: US/09/826,752
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 06/396,001
; PRIOR FILING DATE: 1995-02-28
; PRIOR APPLICATION NUMBER: PCT/US94/09351
; PRIOR FILING DATE: 1994-08-15
; PRIOR APPLICATION NUMBER: US 08/107,408
; PRIOR FILING DATE: 1993-08-16
; PRIOR APPLICATION NUMBER: US 09/322,433
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-826-752-6

Query Match          1.2%; Score 15; DB 9; Length 888;
```



```

; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 298
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-09-801-368-298

Query Match      1.1%; Score 14; DB 10; Length 1770;
Best Local Similarity 100.0%; Pred. No. 0.00084;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      742 NNNNNNNNNNNNNN 755
      |||
Db      345 NNNNNNNNNNNNNN 358

RESULT 27
US-10-032-585-7237
; Sequence 7237, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, JIang
; APPLICANT: Charles, Boone
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7237
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Candida albicans
; US-10-032-585-7237

Query Match      1.0%; Score 13; DB 12; Length 397;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      742 NNNNNNNNNNNNNN 754
      |||
Db      177 NNNNNNNNNNNNNN 189

RESULT 28
US-10-032-585-7930
; Sequence 7930, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, JIang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
```

```

; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7930
; LENGTH: 941
; TYPE: PRT
; ORGANISM: Candida albicans
; US-10-032-585-7930

Query Match      1.0%; Score 13; DB 12; Length 941;
Best Local Similarity 100.0%; Pred. No. 0.0041;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      741 TNNNNNNNNNNNNN 753
      |||
Db      654 TNNNNNNNNNNNNN 666

RESULT 29
US-09-801-368-134
; Sequence 134, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 134
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-09-801-368-134

Query Match      1.0%; Score 13; DB 10; Length 1151;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      502 NNNNNNNNNNNNS 514
      |||
Db      38 NNNNNNNNNNNNS 50

RESULT 30
US-10-060-019-31
; Sequence 31, Application US/10060019
; Publication No. US20030003564A1
; GENERAL INFORMATION:
; APPLICANT: Tyers, Mike
; APPLICANT: Williams, Andrew
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN
; FILE REFERENCE: 11757.10USU1
; CURRENT APPLICATION NUMBER: US/10/060,019
```



```

; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US/09/177,165
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/092,443
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: 60/063,254
; PRIOR FILING DATE: 1997-10-24
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-10-060-019-31

Query Match
Best Local Similarity 100.0%; Score 13; DB 15; Length 1151;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 502 NNNNNNNNNNNNS 514
Db 38 NNNNNNNNNNNNS 50

RESULT 31
US-10-032-585-7034
; Sequence 7034, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7034
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Candida albicans
; US-10-032-585-7034

Query Match
Best Local Similarity 100.0%; Score 12; DB 12; Length 387;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 499 INNNNNNNNNN 510
Db 238 INNNNNNNNNN 249

RESULT 32
US-09-925-300-1074
; Sequence 1074, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1074
; LENGTH: 410
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (14)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (21)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (177)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (248)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (300)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (372)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-300-1074

Query Match
Best Local Similarity 100.0%; Score 12; DB 10; Length 410;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 CFAYGQTSGSKT 102
Db 143 CFAYGQTSGSKT 154

RESULT 33
US-10-032-585-7788
; Sequence 7788, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7788
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Candida albicans
; US-10-032-585-7788

Query Match
Best Local Similarity 100.0%; Score 12; DB 12; Length 511;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 457 NNNNNNNNNNNN 468
Db 50 NNNNNNNNNNNN 61

RESULT 34
US-09-849-602-20
; Sequence 20, Application US/09849602
; Publication No. US20030165834A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Old, Lloyd J.
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Chen, Yao-Tseng
; TITLE OF INVENTION: Colon Cancer Antigen Panel
; FILE REFERENCE: L0461/7105(JRV)
```



```
; CURRENT APPLICATION NUMBER: US/09/849,602
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 725
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-602-20

Query Match
Best Local Similarity 100.0%; Score 12; DB 12; Length 725;
Pred. No. 0.028;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 CFAAGTGTGGCT 102
DB 344 CFAAGTGTGGCT 355

RESULT 35
US-09-924-154-17
; Sequence 17, Application US/09924154
; Patent No. US20020127241A1
; GENERAL INFORMATION:
; APPLICANT: Narum, David L.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
; FILE REFERENCE: 05213-0465 43170-262105
; CURRENT APPLICATION NUMBER: US/09/924,154
; PRIOR FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: US 60/223,525
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1501
; TYPE: PRT
; ORGANISM: Mammalian
US-09-924-154-17

Query Match
Best Local Similarity 100.0%; Score 12; DB 10; Length 1501;
Pred. No. 0.052;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 503 NNNNNNNNNNS 514
DB 46 NNNNNNNNNNS 57

RESULT 36
US-10-106-698-4938
; Sequence 4938, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4938
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
```

```
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-4938

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 160;
Pred. No. 0.061;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 AYGQTSGKTY 103
DB 89 AYGQTSGKTY 99

RESULT 37
US-09-934-455-462
; Sequence 462, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddie, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pineda, Omaira
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: MBI-0022
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: MBI-0023
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 462
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-934-455-462

Query Match
Best Local Similarity 100.0%; Score 11; DB 11; Length 250;
Pred. No. 0.091;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 502 NNNNNNNNNNN 512
DB 4 NNNNNNNNNNN 14

RESULT 38
US-10-286-264-36
; Sequence 36, Application US/10286264
; Publication No. US20030093837A1
; GENERAL INFORMATION:
; APPLICANT: Keddie, James
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Zhang, James
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Pineda, Omaira
; APPLICANT: Heard, Jacqueline
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Adam, Luc
; APPLICANT: Broun, Pierre
; APPLICANT: Reuber, Lynne
```



```
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Samaha, Raymond
; TITLE OF INVENTION: POLYNUCLEOTIDES FOR SEED TRAIT ALTERATION
; FILE REFERENCE: MBI-008
; CURRENT APPLICATION NUMBER: US/10/286,264
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1362
US-10-286-264-36
```

```
Query Match          0.9%; Score 11; DB 15; Length 298;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 504 NNNNNNNNS 514
DB 204 NNNNNNNNS 214
```

```
RESULT 39
US-10-278-536-188
; Sequence 188, Application US/10278536
; Publication No. US20030131386A1
; GENERAL INFORMATION:
; APPLICANT: Samaha, Raymond
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Keddle, James
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Adam, Luc
; APPLICANT: Broun, Pierre
; FILE REFERENCE: MBI-011
; TITLE OF INVENTION: STRESS-INDUCED POLYNUCLEOTIDES
; CURRENT APPLICATION NUMBER: US/10/278,536
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 188
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1362
US-10-278-536-188
```

```
Query Match          0.9%; Score 11; DB 16; Length 298;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 504 NNNNNNNNS 514
DB 204 NNNNNNNNS 214
```

```
RESULT 40
US-09-991-496-134
; Sequence 134, Application US/09991496
; Patent No. US20020169285A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, David C.
; APPLICANT: Skelley, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C9
; CURRENT APPLICATION NUMBER: US/09/991,496
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Leishmania major and chagasi
US-09-991-496-134
```

```
Query Match          0.9%; Score 11; DB 10; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 92 FAYGOTSGKT 102
DB 256 FAYGOTSGKT 266
```

```
Search completed: October 2, 2003, 16:45:20
Job time : 102 secs
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 2, 2003, 16:07:05 ; Search time 29 Seconds

(without alignments)
2088.634 Million cell updates/sec

Title: US-10-006-780-2

Perfect score: 6842
Sequence: 1 MNSKIKVVKRPLSELEK.....KLIQDNKSNMDNNHHK 1288

Scoring table: BLAST62

Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	761.5	11.1	730	1 KCM1_XENLA	Q91636 xenopus lae
2	752.5	11.0	725	1 KF2C_HUMAN	Q99661 homo sapien
3	747.5	10.9	671	1 KF2C_MACFA	Q95161 macaca fasc
4	735.5	10.7	679	1 KIF2_HUMAN	O00139 homo sapien
5	735	10.7	718	1 KF2C_CRIGR	P70096 cricetus
6	732	10.7	721	1 KF2C_MOUSE	Q92268 mus musculu
7	728	10.6	716	1 KIF2_MOUSE	P28740 mus musculu
8	722	10.6	671	1 KF2C_RAT	Q62909 ratius norv
9	719.5	10.5	682	1 KIF2_XENLA	Q91637 xenopus lae
10	597	8.4	537	1 ARP_PLAFA	Q94931 plasmodium
11	576	8.4	624	1 DSK1_CYLFU	Q94931 plasmodium
12	561.5	8.2	1111	1 KIF1_YEAST	P28742 saccharomyc
13	560.5	8.2	2339	1 RPP1_PLAFA	P27665 plasmodium
14	543	7.9	1956	1 ATX1_PLAFA	Q04956 plasmodium
15	538.5	7.9	1085	1 CUT7_SCHPO	P24339 schistosach
16	520.5	7.6	1056	1 K125_ARATH	P82266 arabidopsis
17	513	7.5	1006	1 K125_TOBAC	O23896 nicotiana t
18	500.5	7.3	967	1 KINH_LOJPE	P21613 loligo peal
19	500	7.3	1184	1 BIMC_EHMTI	P17100 emericella
20	482.5	7.1	2663	1 CENE_HUMAN	Q02224 homo sapien
21	477	7.0	784	1 KL68_DROME	P46867 drosophila
22	472.5	6.9	747	1 KF3B_HUMAN	O15066 homo sapien
23	472.5	6.9	747	1 KF3B_MOUSE	O61771 mus musculu
24	465.5	6.8	554	1 KLUP3_SCHPO	Q9560 schistosach
25	465.5	6.8	1585	1 P3K3_DICDI	P54675 dictyosteli
26	463	6.8	1038	1 C1N8_YEAST	P27895 saccharomyc
27	460	6.7	702	1 KF3A_HUMAN	Q94464 homo sapien
28	459	6.7	1066	1 KL61_DROME	P46863 drosophila
29	453	6.6	963	1 KIF3_HUMAN	P33176 homo sapien
30	452.5	6.6	805	1 KIP3_YEAST	P33066 saccharomyc
31	449	6.6	883	1 KLUP3_SCHPO	O14343 schistosach
32	447.5	6.5	1057	1 KFI1_HUMAN	P52732 homo sapien
33	447	6.5	963	1 KINH_MOUSE	O61768 mus musculu

34	446	6.5	1648	1 KF14_HUMAN	Q15058 homo sapien
35	445.5	6.5	701	1 KF3A_MOUSE	P28741 mus musculu
36	443.5	6.5	956	1 KF5C_MOUSE	P28738 mus musculu
37	442.5	6.5	1236	1 KF4A_XENLA	O91784 xenopus lae
38	439.5	6.4	1067	1 EG52_XENLA	O91783 xenopus lae
39	439	6.4	957	1 KF5C_HUMAN	O60282 homo sapien
40	439	6.4	975	1 KINH_DROME	P17210 drosophila
41	437.5	6.4	1231	1 KF4A_MOUSE	P33174 mus musculu
42	435	6.4	786	1 FL10_CHLRE	P46669 chlamydomon
43	435	6.4	989	1 PTP3_DICDI	P54637 dictyosteli
44	434	6.3	2452	1 RPP1_PLAFA	P14248 plasmodium
45	433.5	6.3	742	1 K121_STRPU	P46871 strongyloce

ALIGNMENTS

RESULT 1
ID KCM1_XENLA STANDARD; PRT; 730 AA.
AC Q91636;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Kinesin central motor 1 (XKCM1).
GN KCM1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=96140638; PubMed=8548824;
RA Walczak C.E., Mitchison T.J., Desai A.;
RT "XKCM1: a Xenopus kinesin-related protein that regulates microtubule
RT dynamics during mitotic spindle assembly.";
RL Cell 84:37-47(1996).
RN [2]
RP REVISION TO 516.
RC TISSUE=Ovary;
RA Walczak C.E.;
RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: REGULATES MICROTUBULE DYNAMICS DURING MITOTIC SPINDLE
CC ASSEMBLY. REQUIRED FOR BOTH ESTABLISHMENT AND MAINTENANCE OF
CC MITOTIC SPINDLES. MAY ACT DIRECTLY ON MICROTUBULES, CAUSING A
CC DESTABILIZATION AND EVENTUAL DEPOLYMERIZATION OF THE MICROTUBULE.
CC -!- SUBUNIT: Homodimer (potential).
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. MCAK/KIF2
CC SUBFAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U36485; AAC59743.2; -.
CC HSSP: P17119; 3KAR
DR InterPro: IPR001752; kinesin_motor.
DR Pfam: PF00225; kinesin_1.
DR PRINTS: PR00380; KINESINHEAVY.
DR SMART: SM00129; KISC; 1.
DR PROSITE: PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE: PS50067; KINESIN MOTOR DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
FT DOMAIN 1 256 GLOBULAR (POTENTIAL).
FT DOMAIN 257 598 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 599 730 COILED COIL (POTENTIAL).

FT NP BIND 352 359 ATP (POTENTIAL).
 SQ SEQUENCE 730 AA; 82585 MW; 25031C187E491523 CRC64;
 Query Match 11.1%; Score 761.5; DB 1; Length 730;
 Best Local Similarity 42.1%; Pctd. No. 1.3e-26;
 Matches 175; Conservative 85; Mismatches 129; Indels 27; Gaps 12;

4 KIKVVRKRPSELEKKKSDIIVKNNCTYIDPEYKVDMTYIERHEIYDKVDD 63
 262 RLCVVRKRPSELEKKKSDIIVKNNCTYIDPEYKVDMTYIERHEIYDKVDD 321
 64 TVDNFTYENTIKPLIIDLKNGCVSCFAYGQTSKGYTLTGS-QPYGSDPFGIRFOY 122
 322 TATNEVVVRFPARPLVQSFEG-KATCFAYGQTSKGYTLTGS-QPYGSDPFGIRFOY 380
 123 AAGDIFTLNTYDKNTK---GIFISFEYICGLYLLQKRVAALENGKAVVVKD 178
 381 ASRDVFL---DQRYKHLDLVFPTEIYNGKVPDLNKKTKRLLEDAKQEVYVG 437
 179 LKILRVLRKELIKMID-GVLLRKIGVNSQNDSSRSRSHATINDKDKINKTSL-GKIA 236
 438 LLEKOVISADD-VFMMETIGSACRTSGQTPANTSSSRSHACIQLTL---RQSKLHGKFS 493
 237 FIDLAGSERGADTVSQNKQOTGDGAINRSLALKECIRAMDSDKNI-PFRDSELTXYLR 296
 494 LVDLAAGNERGVDTASADRITRMEGALNRSLLALKECIRALGQNKSHTPRESKLTQILR 553
 297 DIPVCK-SKSMINIANISPTISCEQTLNTRSSVYKFKKSTINEEDNTIRISIL 355
 554 DSFIENSRCTCIAMLSFGFNSCEYTLNTRVADYVKELSPONA-----ETNDNQLQME 607
 356 DSKSEKNASSIENNVVINSNHLNNNNKINRGKINDKIEENNIKAKSPDKPE 411
 608 DSGSH---ASIEGLQLODDFLKDEBLSTAN--SFQALNRVGLBKAVDLE 658

Db 608 DSGSH---ASIEGLQLODDFLKDEBLSTAN--SFQALNRVGLBKAVDLE 658

RESULT 2
 KF2C HUMAN STANDARD; PRT; 725 AA.
 ID_KF2C_HUMAN STAN: 725 AA.
 AC Q99661; Q96C18; Q96HB8; Q96WV8;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 28-FEB-2003 (Rel. 41; Last sequence update)
 DT 15-SEP-2003 (Rel. 42; Last annotation update)
 DE Kinesin-like protein KIF2C (Mitotic centromere-associated kinesin)
 DE (MCAK) (Kinesin-like protein 6).
 GN KIF2C OR KNSL6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NX NCBI_Taxid=9606;
 RX [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=T-cell;
 RC MEDLINE=98094213; PubMed=9434124;
 RA Kim I.-G., Jun D.-Y., Sohn U., Kim Y.H.;
 RT "Cloning and expression of human mitotic centromere-associated kinesin
 gene";
 RL Biochim. Biophys. Acta 1359:181-186 (1997).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Testis;
 RA Cheng L.J.;
 RT "Cloning of human testis specific mitotic centromere-associated
 kinesin";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Kidney; and Uterus;
 RC MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan G.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stedilton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toehlyki S., Carninci P., Prange C.,
 RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.U., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wozley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalek A., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -1- FUNCTION: Present throughout the cell cycle, associates with the
 CC centromeres at early prophase, and remains associated with the
 CC centromere until after telophase (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q99661-1; Sequence=Displayed;
 CC Name=2; Synonyms=TS-MCAK;
 CC IsoId=Q99661-2; Sequence=VSP 002866;
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THYMUS AND TESTIS,
 CC AT LOW LEVELS IN SMALL INTESTINE, THE MUCOSAL LINING OF COLON, AND
 CC PLACENTA, AND AT VERY LOW LEVELS IN SPLEEN AND OVARY; EXPRESSION
 CC IS NOT DETECTED IN PROSTATE, PERIPHERAL BLOOD LEUKOCYTES, HEART,
 CC BRAIN, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY, OR PANCREAS.
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. MCAK/KIF2
 CC SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

 CC EMBL; U63743; AAC27660.1; -;
 CC EMBL; AY026505; AAK20168.1; -;
 CC EMBL; BC014924; AAH14924.1; -;
 CC EMBL; BC008764; AAH08764.1; -;
 CC HSSP; P17119; 3KAR.
 CC DR GENE; HGNC:6393; KIF2C.
 CC MIM; 604538; -;
 CC GO; GO:0005871; C:kinesin complex; TAS.
 CC GO; GO:0003777; F:microtubule motor activity; TAS.
 CC GO; GO:0008283; P:cell proliferation; TAS.
 CC GO; GO:0007002; P:centromere binding; TAS.
 CC GO; GO:0007067; P:mitosis; TAS.
 CC InterPro; IPR001752; Kinesin_motor.
 CC Pfam; PF00225; kinesin; 1.
 CC PRINTS; PR00380; KINESINHEAVY.
 CC SMART; SM00129; KISC; 1.
 CC PROSITE; PS00411; KINESIN MOTOR DOMAIN; 1.
 CC PROSITE; PS50067; KINESIN MOTOR DOMAIN; 1.
 CC Motor protein; Microtubules; ATP-binding; Coiled coil;
 CC Nuclear protein; Alternative splicing.
 CC DOMAIN 1 254 GLOBULAR (POTENTIAL).
 CC DOMAIN 255 518 KINESIN-MOTOR (BY SIMILARITY).
 CC DOMAIN 618 658 COILED COIL (POTENTIAL).
 CC NP_BIND 348 355 ATP (POTENTIAL).
 CC DOMAIN 415 418 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 CC VASPLIC 1 55 MAMMOSLOARPEPLAIGKISNGLISAVRYVNLKSCV
 CC SVWAEKATKGE -> M (in isoform 2).
 CC /ftid=vsp_002866;
 CC I -> L (IN REF. 1 AND 2).
 CC CONFLICT 449 449 R -> P (IN REF. 3; AAH08764).
 CC SEQUENCE 725 AA; 81312 MW; 5BDECC13AB4B55C CRC64;


```

Query Match      11.0%; Score 752.5; DB 1; Length 725;
Best Local Similarity 42.1%; Pred. No. 3,1e-26;
Matches 174; Conservative 75; Mismatches 143; Indels 21; Gaps 10.

OY      4 KIVVVRKRPISLEKKKKSDIITVKNNTLYIDEPKYDMTYIERHEFIYDKVFD 63
Db      258 RIVCVRRKRPINKQELAKKEIDIVISIPSKCLLVHPEKLVLTLYLEQACFPFAFE 317
OY      64 TVDNFTYVNTIKPLIIDLYENGCCVCSGAVQOTSQGYTMUGS-QPYGSDTFGICFY 122
Db      318 TASNNVVRFTRAPLVQTIIFSGG-KATCAVQOTSQGYTHMGULSGKQVNASGITYAM 376
OY      123 AAGDIETPLN--IYDKMTYKGFIFISFEYICGLVDLQKRMVALENGKEVAVKDL 180
Db      377 ASRDVFLKNOQCYRKGLG-YYVFFFEIYNGKLPDLNKKAKKLVLEDGQQQVYVGLQ 435
OY      181 ILRVLTKEELIKMID-GVLAKTIQVNSQNDSSSHALINIDLKQINKNTSLGKIATFD 239
Db      436 EHLVNSAD-VIKMIDWSACITSGQTRANSSSHACFOILN--AKGRMGKFSIYD 492
OY      240 LAGSERGADTVSONKQOTQDGANIRSLALKECRAMPDSQKNIIPRDSSELYLRDIF 299
Db      493 LAGNRRGADTSSADQTMEGEINIKSLALKECIRALGQNKANHPFRESKLTQVLRDSF 552
OY      300 VGR-SKSIIMANISPTISCCQETLNTLRYSFVNFKKSTCINEEDDTNTERISILDSK 358
Db      553 IGENSRITMIAITISGISISCEYTLNTLTLYADRVKELSPHSGPSGQ-----LIQME 603
OY      359 GSENVASSIEENVVIKSHLNSNNNNKINRGKINDIERNNNILNKRSPDKPRE 411
Db      604 TEENHACS--NGALIPGULSKKEEELSOMSFENAMQIIRLEKXANBELKE 654

```

```

RESULT 3
KF2C_MACFA
ID_KF2C_MACFA      STANDARD;      PRT;      671 AA.
AC_Q95DF1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE kinesin-11-like protein KIF2C (Mitotic centromere-associated kinesin)
DE (MCAK).
GN KIF2C.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RX (1)
RN SEQUENCE FROM N.A.
RP TISSUE=Testis;
RC Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
RA Terao K., Sugano S.;
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
RT libraries.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Present throughout the cell cycle, associates with
CC centromeres at early prophase, and remains associated with the
CC centromere until after telophase (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. MCAK/KIF2
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL; AB072747; BAB69716.1; -

```

Query Match	Best Local Similarity	Score	DB 1	Length	DB 2	Length
Matches 173; Conservative 76; Mismatches 143; Indels 21; Gaps 10.	41.9%;	747.5;	DB 1:	671;	DB 2:	671;
DR Interpro: IPR001752; kinesin_motor.						
DR Pfam: PF00225; kinesin_1.						
DR PRINTS: PR00380; KINESINHEAVY.						
DR SMART: SM00129; KISC. 1.						
DR PROSITE: PS00411; KINESIN MOTOR DOMAIN1. 1.						
DR PROSITE: PS50067; KINESIN MOTOR DOMAIN2. 1.						
KV Motor protein; Microtubules; ATP-binding; Coiled coil;						
KV Nucleic protein.						
FT DOMAIN 1. 200						
FT DOMAIN 201						
FT DOMAIN 564						
FT NP_BIND 294						
SO SEQUENCE 671 AA; 75639 MW; 4EA8399E8B6BA46 CRC64;						
Query Match	10.9%;	Score 747.5;	DB 1:	Length 671;	DB 2:	Length 671;
Best Local Similarity	41.9%;	Pred. No. 4.7e-26;				
Matches 173; Conservative 76; Mismatches 143; Indels 21; Gaps 10.						
DR Interpro: IPR001752; kinesin_motor.						
DR Pfam: PF00225; kinesin_1.						
DR PRINTS: PR00380; KINESINHEAVY.						
DR SMART: SM00129; KISC. 1.						
DR PROSITE: PS00411; KINESIN MOTOR DOMAIN1. 1.						
DR PROSITE: PS50067; KINESIN MOTOR DOMAIN2. 1.						
KV Motor protein; Microtubules; ATP-binding; Coiled coil;						
KV Nucleic protein.						
FT DOMAIN 1. 200						
FT DOMAIN 201						
FT DOMAIN 564						
FT NP_BIND 294						
SO SEQUENCE 671 AA; 75639 MW; 4EA8399E8B6BA46 CRC64;						

```

OY      LAGERGADPTYSOKKOTQDGTGNINRSLALKEICIRAMDSPKHNIIPRDSLETVNRIF   299
         |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB       LAGNEREDITSSARQRTRMBGAELNKSLLAKBETCRALGGVKATPFRSKLTVDLRSF   498
                                         :|
OY       300 VGR-SKSMIANTISPTISCCEQTINTLRYSRVNFNNKCSTCINEBDTWTERISTLSK    358
          :::::|||||::|:
DB        IGENSRRCMTATISPGLSSECYTLNTLTRADVRELSPHSQPSEQ-----LTQME     549
                                         |
OY       359 GSEMNASISIENVIKSNHLISNNNNKNINKNGINDKIERNNILKKNSFKDKPRE   411
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db        550 TEEWEACS--NGALIPLGNLSKEBEHELSSOWMSFEAMTQLRELEBRAVEELKE    600
                                     :|
RESULT 4
KIF2_HUMAN                                STANDARD;              PRT;           679 AA.
AC      KIF2_HUMAN                        000139;
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      15-JUL-1998 (Rel. 36, Last annotation update)
DE      Kinesin-like protein KIF2 (Kinesin-2) (HK2).
GN      KIF2 OR KINS2.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=97321046; PubMed=9177777;
RA      Debernard S., Fontanella E., de Gregorio L., Pierotti M.A.,
RD      Delta D.;
RT      "Identification of a novel human kinesin-related gene (HK2) by the
RL      cDNA differential display technique.";
CC      Genomics 42:67-73(1997).
-i- FUNCTION: THE KINESIN FAMILY MAY RECOGNIZE, TRANSPORT AND POSITION
CC      THEIR SPECIFIC CARGOES IN A SINGLE TYPE OF NEURONAL CELL.
CC      -i- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. WCAK/KIF2
SUBFAMILY.
```


CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/announcement/> or send an email to license@ebi.ac.uk).

CC -----

DR EMBL; Y08319; CAA69621.1; -.
DR HSBP; F17119; 3KAR.
DR Genew; HGNC:6318; KIF2.
DR MIM; 602591; -.
DR GO; GO:0004002; F:adenosinetriphosphatase activity; TAS.
DR GO; GO:0003774; F:motor activity; TAS.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin_1.
DR PRINTS; PR00380; KINESIN1HEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
KM Motor protein; Microtubules; ATP-binding; Coiled coil; Neutrone.
FT DOMAIN 1 190 GLOBULAR (POTENTIAL).
FT DOMAIN 191 525 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 525 679 COILED COIL (POTENTIAL).
FT NP BIND 286 293 ATP (BY SIMILARITY).
SQ SEQUENCE 679 AA: 76927 MW; F96A8BBA671BCEB CRC64;

Query Match 10.7%; Score 735.5; DB 1; Length 679;
Best Local Similarity 45.8%; Pred. No. 1.6e-25;
Matches 167; Conservative 62; Mismatches 111; Indels 25; Gaps 9;

QY 4 KIKVYRRRRPSELEKKKKDDITVKNKNTLYIDEPYKYDMTYIRHEIYDVKVDD 63
DB 196 RICVCVRRRPLNKKETQMKDDIVITIPSKDVVMVHPEKQKVDLTRYLENQTFRFYVAFDD 255
QY 64 TVDNFTVVENTIKPLIIDLYENGCVCSGFAYGQTSQGTYYMLGS-QPYGSDTFPGIFQY 122
DB 256 SAPNEMVYRFNAKPLVETIFERNG-MATCFAYQQTSSGKTHYMGDFPSGRKDDCSGIYAL 314
QY 123 AAGDIFTEPLNIYDKDNTK---GIFISFEYIYCGLYDLLOKRYVAALENGKEVYVD 178
DB 315 AARDVFALMK--KENYKKLELYQVATFEIYSGKVPDLNRKTKLRFVLEDKQGVVVG 371
QY 179 LKILRVLTKEELILMKID-CYLIRKICGVNSQNDSSRSHTALINDLKIINKTSLSGIAF 237
DB 372 LQEREVKVED-VLKLIDIGNSCRFSQGTSAVAHSSRHAVPQIIIL--RKGLKGFSL 428
QY 238 IDLASERGADIVSONKQOTQDGANINRSIALKSCICRAMSDKXHIIPRDESLTKVYRD 297
DB 429 IDLAENEGADISSADROTRLEGATINSLSLAKCICIRALGKNNKHTTFFRSKALKVYLRD 488
QY 298 IFVGR-SKSIIMANISPTISCCQOTLNTLRYSRYVNFKNKSTCI-----NEED 345
DB 489 SFIGENSTQMTATISPGMASCENTLNTLRANRYKELVDPDPAAGVYRPIHHHPNQID 548
QY 346 DTNTE 350
DB 549 DLETQ 553

RESULT 5
KIF2C CRIGR STANDARD; PRT; 718 AA.
AC P70056;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE kinesin-like protein KIF2C (Mitotic centromere-associated kinesin)
DE (MCAK) (Kinesin-1like protein 6).
DE KIF2C OR KNSL6.
OS Cricetulus griseus (Chinese hamster).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OJ Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;

```

OC Cricetulus.
OX NCBI_TaxID=10029;
RN
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95122643; PubMed=7822426;
RA Wordeman L., Mitchson T.J.;
RT "Identification and partial characterization of mitotic centromere-
RT associated kinesin, a kinesin-related protein that associates with
RT centromeres during mitosis.";
RL J. Cell Biol. 128:95-105(1995).
RN
RN [2]
RP REVISIONS.
RA Wordeman L.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Present throughout the cell cycle, associates with
CC centromeres at early prophase, and remains associated with the
CC centromere until after telophase.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. MCAK/KIF2
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U11790; AAB17358.2; -.
DR HSSP; P17119; 3KAR.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin_1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; Kisc; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
DR Motor protein; Microtubules; ATP-binding; Coiled coil;
KM Nuclear protein.
KT DOMAIN 1 248 GLOBULAR (POTENTIAL).
FT FT 249 512 KINESIN-MOTOR.
FT DOMAIN 613 651 COILED COIL (POTENTIAL).
FT DOMAIN 689 716 COILED COIL (POTENTIAL).
FT NP_BIND 342 349 ATP (POTENTIAL).
FT FT 409 412 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 718 AA; 80918 MW; 1AABDBCB6AD11B2 CXC64;

Query Match 10.7%; Score 735; DB 1; Length 718;
Best Local Similarity 40.9%; Pred. No. 1.8e-25;
Matches 172; Conservative 75; Mismatches 136; Indels 38; Gaps 9;

OY 4 KIKVVRKRPISLEKKKSDIITVKNKNTLYIDEPRIKYDKMTKYLIRHFIYDKVDD 63
DB 252 RICVCRKRPINKOELAKKEIDIVSVSKCLFVHEPKLKYDLTKYLENQAFCDFAPDE 311
OY 64 TVDNFTYENTIKPLIIDLNGCVCSCEFAAGTGGSGKTYMTLS-QYGGSDPIGFIQY 122
DB 312 TASNEVYRFAARPLVOTIFEGG-KATFEAIGGTSGKTHMGDLGSKSONTSGIYAM 370
OY 123 AAGDIPTF-----LNIYDKMTKGIFISFEYIECGKLYDLLQKRMVAALENGKKEV 174
DB 371 ASRDVFLKSPRRIRNLN-----EYVYVFFELYNKRVFDLNRKAKKAVLDBSKQOV 423
OY 175 VKDKILRLVLTKEELILKMLDGVLRKIGVNSONDESSSHAILNIDLKINKNTSLGK 234
DB 424 QVVGLEQEVLVNACDDVIMGLNMGSACTRSQGTFAANSNSRSHACFOILLR--AKGRLEHGK 481
OY 235 IAFIDLASSEGGADTVSONKQOTQDNGANINSLAKKECTIAMSDDKNHIPPROSELTKV 294
DB 482 FSLVDLAINEGADITSSADRTIRMEGAIRINSLALKECICRALGQNKAKHTPPRESKLTQV 541
OY 295 LRDIFVVGK-SKSIIMANISPTISCEQRLNLTARYSRKFNKNTCTCINEEDDTWTERIS 353

```



```

Db 542 LRDSFGNSRTNCMIAMISPGISCEXTLNTLRVADRKELSPHSIGSGEOP----- 593
Qy 354 ILDSKSGENNAASSTEN---VVIKSNHLLSNNNKINRGINDKIERNNILKNKSPDKPR 410
Db 594 -IQMTEEMASSNGTSLAVNFKEEELSSOMSS-----FNEAMSQIRELENAMEELR 646
Qy 411 E 411
Db 647 E 647

RESULT 6
KF2C MOUSE STANDARD; PRT; 721 AA.
ID KF2C MOUSE
AC 092258;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kinesin-like protein KIF2C (Mitotic centromere-associated kinesin)
DE (MCAK).
GN KIF2C.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uscdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -I- FUNCTION: Present throughout the cell cycle, associates with the
CC centromeres at early prophase, and remains associated with the
CC centromere until after telophase (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
CC -I- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. MCAK/KIF2
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: BC006841, AA006841.1; -
DR MGI: MGI:1921054; KIF2C.
DR InterPro: IPR001752; kinesin_motor.
DR Pfam: PF00225; kinesin_1.
DR PRINTS: PR00380; KINESINHEAVY.
DR SMART: SM00129; KISC; 1.
DR PROSITE: PS00441; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE: PS50067; KINESIN MOTOR DOMAIN2; 1.
DR Motor protein; Microtubules; ATP-binding; Coiled coil;
KW Nuclear protein.

```

```

FT DOMAIN 1 250 GLOBULAR (POTENTIAL).
FT DOMAIN 251 514 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 614 652 COILED COIL (POTENTIAL).
FT NP BIND 348 355 ATP (POTENTIAL).
SQ SEQUENCE 721 AA; 81085 MW; 92FFBFFA05B7E35 CRC44;

Query Match 10.7%; Score 732; DB 1; Length 721;
Best Local Similarity 41.2%; Pred. No. 2,4e-25;
Matches 170; Conservative 78; Mismatches 141; Indels 24; Gaps 10;

Qy 4 KIVVVRKRLPSELEKKKDSIDITVKNNTLYIDPRYVDMTKYIERHIEIVDKFDD 63
Db 254 RICVCVRKRLPKNQELAKKEIDIVSPSKCLLVHBPVKLVTKYLENQAFDFPAFDE 313
Qy 64 TVNFTYVNTTPELTIIDYENGCVCSFPAVGQTSKTYTMS- QPVGQSPGPGFY 122
Db 314 TASNEVYRTARPLVQTTFEGG-KATCPAYGQTSKTHWGDDLSGSKNSKGIYAM 372
Qy 123 AAGDIFFTLN--IYDKDNTKGFISFYEIYCGLYDLQKRMVAALENGKEVWVDLK 180
Db 373 ASRDVFLKNQPRYRNLANLE-VYTFEITNGKVPFDLNNKATLRYLEDSDRQVGVGLQ 431
Qy 181 ILRVLTKEELIMID-GVLLRKIGVNSQNDSSRSHALINDKDKNTSLGKIAFID 239
Db 432 EYLVTCADD-VIKIMNGSACRTSGQTFANSSSRSHACFQILLR--TKGRHLGKESLYD 488
Qy 240 LAGSERGADTVSONKQOTDGDANINRSLALKECIRAMSDKHIFRDELTKVLRDIF 299
Db 489 LAENRGADTVSSADRTREGAINGINSLALKECIRALGONKATPRESKLTQVLRDSF 548
Qy 300 VGR-SKSIMANISPTISCEQTLNTRYSSRYKFNKNSCTCINEEDDTTERISILDSK 358
Db 549 IGENSTCMQAMISPSISGEYTLNLRVADRKELSPHSIGSGEOPVQMETEMASSN 608
Qy 359 GSEMAASSIENVVYKSNHLLSNNNKINRGINDKIERNNILKNKSPDKPR 411
Db 609 GTSITGNEBEL---SSOMSS-----FNEAMTOIRELEERALELRE 647

RESULT 7
KF2C MOUSE STANDARD; PRT; 716 AA.
ID KF2C MOUSE
AC P28740;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kinesin-like protein KIF2.
GN KIF2 OR KIF2A OR KNS2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93077686; PubMed=1447303;
RA Aizawa H., Sekine Y., Takemura R., Zhang Z., Nangaku M.,
RA Hirokawa N.;
RT "Kinesin family in murine central nervous system."
RL J. Cell Biol. 119:1287-1296(1992).
CC -I- FUNCTION: THE KINESIN FAMILY MAY RECOGNIZE, TRANSPORT AND POSITION
CC THEIR SPECIFIC CARGOS IN A SINGLE TYPE OF NEURONAL CELL.
CC -I- TISSUE SPECIFICITY: EXPRESSED AMONG VARIOUS TISSUES UNOUTOUSLY;
CC ITS EXPRESSION LEVEL IN BRAIN DECREASES WITH DEVELOPMENT.
CC -I- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. MCAK/KIF2
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```


or send an email to license@sb-sib.ch.

CC EMBL; D12644; BAA02165.1; -
 DR PIR; A44259; A44259.
 DR HSSP; P17119; 3KAR.
 DR MGD; MGI.108390; Klf2a.
 DR InterPro; IPR001752; kinesin_motor.
 DR Pfam; PF00225; kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
 DR Motor protein; Microtubules; ATP-binding; Coiled coil; Neurone.
 FT DOMAIN 1 189 GLOBULAR (POTENTIAL).
 FT DOMAIN 190 539 KINESIN-MOTOR (BY SIMILARITY).
 FT DOMAIN 540 716 COILED COIL (POTENTIAL).
 FT NP BIND 285 292 ATP (BY SIMILARITY).
 SQ SEQUENCE 716 AA; 80946 MW; 6653D01F00138A1B CRC64;

Query Match 10.6%; Score 728; DB 1; Length 716;
 Best Local Similarity 47.6%; Pred. No. 3.6e-25;
 Matches 161; Conservative 61; Mismatches 102; Indels 14; Gaps 8;

QY 4 KIKVVRKRPPLSELEKKKKSDITTVKNNCTLYIDEPKYVDMTKYIERHFEIVDKVED 63
 195 RICVVRKRPPLNKKEKQKDDIVITIPSKDVVMVHEPQKVDLTRELENQTFRFPYARD 254
 QY 64 TVDNFTYENTIKPILIDLYENGCVCSFAYGQTSKTYTMLS-QPYGSDTPGIFQY 122
 255 SAPENMYRFTARPLVQTFIFERG-KATCFAYGQTSKTYTMGDLSGKSNAGKIYAL 313
 QY 123 AAGDIFTPLNTYDKONTK---GIFISFEIYCGLYDLLQRRKVAALENGKEVVKD 178
 314 AARDVFLMLK--KENYKLELOVATFEIYSGVFPPLNRKTKTLRLLEDKQOVVVG 370
 QY 179 LKILRLTKEELIKMID-GVLLRKIGVNSQNDSSRSALINDLKDINKNTSLGKIAF 237
 371 LOERVKVED-VLKLIDIGNSCRISGQTSANVHSSRAVQIILR--RKGKLHGKPSL 427
 QY 238 IDLAGSEGADTVSONKQOTQDGNINRSILALKECIKRAMSDKNHIFPDSLTLYARD 297
 428 IDLNGERGADTVSADRQTRLEGAINSLRLKCEICIRALGKHPHTFRASKLQVARD 487
 QY 298 IFVGK-SKSIIMANISPTISCEQTLNTRYSSRYKNF 334
 488 SFIENSRCTMIAITISPGMASCENTLNTLRANRYKEF 525

RESULT 8
 KFP2C RAT STANDARD; PRT; 671 AA.
 AC Q62909
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Kinesin-like protein KIF2C (Mitotic centromere-associated kinesin)
 DE (MCAK) (Kinesin-related protein 2).
 GN KIF2C OR KRP2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Testis;
 RX MEDLINE=96228687; PubMed=8688559;
 RA Sperry A.O., Zhao L.-P.;
 RT "Kinesin-related proteins in the mammalian testes: candidate motors
 for meiosis and morphogenesis";
 RL Mol. Biol. Cell 7:289-305(1996).
 RN [2]
 RP REVISIONS.
 RA Sperry A.O.;

RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 CC -! FUNCTION: Present throughout the cell cycle, associated with the
 CC centromeres at early prophase, and remains associated with the
 CC centromere until after telophase (by similarity). Active in
 CC meiosis.
 CC -! SUBCELLULAR LOCATION: Cytoplasmic and nuclear (by similarity).
 CC -! TISSUE SPECIFICITY: TESTIS. LOCALIZED TO THE MEIOTICALLY ACTIVE
 CC CELLS OF THE SEMINIFEROUS EPITHELIA IN THE TESTIS.
 CC -! SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. MCAK/KIF2
 CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@sb-sib.ch).

CC EMBL; U44979; AAC53528.1; -
 DR PIR; T10755; T10755.
 DR HSSP; P17119; 3KAR.
 DR InterPro; IPR001752; kinesin_motor.
 DR Pfam; PF00225; kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
 DR Motor protein; Microtubules; ATP-binding; Coiled coil;
 KW Nuclear protein; Meiosis.
 FT DOMAIN 1 200 GLOBULAR (POTENTIAL).
 FT DOMAIN 201 464 KINESIN-MOTOR.
 FT DOMAIN 564 669 COILED COIL (POTENTIAL).
 FT NP BIND 294 301 ATP (POTENTIAL).
 SQ SEQUENCE 671 AA; 75661 MW; F2B5459878B98BE CRC64;

Query Match 10.6%; Score 722; DB 1; Length 671;
 Best Local Similarity 41.2%; Pred. No. 6.1e-25;
 Matches 171; Conservative 78; Mismatches 138; Indels 28; Gaps 11;

QY 4 KIKVVRKRPPLSELEKKKKSDITTVKNNCTLYIDEPKYVDMTKYIERHFEIVDKVED 63
 204 RICVVRKRPPLNKKEKQKDDIVITIPSKDVVMVHEPQKVDLTRELENQTFRFPYARD 263
 QY 64 TVDNFTYENTIKPILIDLYENGCVCSFAYGQTSKTYTMLS-QPYGSDTPGIFQY 122
 264 TASNEMVYRFTARPLVQTFIFERG-KATCFAYGQTSKTYTMGDLSGKSNAGKIYAM 322
 QY 123 AAGDIFTPLN--IYDKONTKGIIFSFEIYCGLYDLLQRRKVAALENGKEVVKD 180
 323 ASRVVFLKQNPFRYSINLE-VVYTFEIVNGKVFELNKKAKRLVLEDSKQOVVGLQ 381
 QY 181 ILRLVLTKEELIKMID-GVLLRKIGVNSQNDSSRSALINDLKDINKNTSLGKIAF 239
 382 EYLVTCADD-VIKKINNGSACRTSGQTFANSSRSASHACFOILR--AKGRHGKFSVD 438
 QY 240 LAGSEGADTVSONKQOTQDGNINRSILALKECIKRAMSDKNHIFPDSLTLYARD 299
 439 LAGNERGADTVSADRQTRLEGAINSLRLKCEICIRALGKHPHTFRASKLQVARD 498
 QY 300 VGG-SKSIIMANISPTISCEQTLNTRYSSRYKNFKNKSTCINEEDTNTERTISLDSK 358
 499 IGENSRCTMIMISPGISSCEYTLNTRYADRVKELSPHSGPSGEO-----AVQME 549
 QY 359 GSEMASSIEENVVYKSNHLSNNNNKIN--RGKINDKIENNLLKNSPKPRE 411
 550 TEEDVASS-----HGASLTGNNEEBLSQMSSEFNAMTQIRLEERAMELE 597
 DB

RESULT 9
 KIF2_XENLA STANDARD; PRT; 682 AA.
 AC Q91637;

Y	298	FVAK-SKSI	MIANISPTIS	CC	EOPLINTLR	SRSPYCNK	STC	340								
Db	491	SFGNSRT	CMATIS	SP	GMASCE	INTLTRYAN	RKEL-DBSRC	533								
RESULT 10																
ID	ARP_PLAFA	STANDARD	PRT	537	AA											
AC	P04931															
DT	13-AUG-1987	(Rel. 05, Created)														
DT	01-NOV-1995	(Rel. 32, Last sequence update)														
DT	01-NOV-1995	(Rel. 32, Last annotation update)														
DE	Asparagine-rich protein (AG319) (ARP) (Fragment).															
OS	Plasmodium falciparum.															
CC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.															
OX	NCBI_TaxID=5833;															
RN	[1]															
RP	SEQUENCE FROM N.A.															
RA	MEDLINE=66176787; PubMed=2421257;															
RA	Stehli H.-D., Bianco A.E., Crewther P.E., Burkot T., Coppel R.L.,															
RT	Brown G.V., Anders R.F., Kemp D.J.;															
RT	"An asparagine-rich protein from blood stages of Plasmodium															
RL	falciparum shares determinants with sporozoites.";															
CC	Nucleic Acids Res. 14:3089-3102(1986).															
CC	-----															
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration															
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation.															
CC	the European Bioinformatics Institute. There are no restrictions on its use															
CC	by non-profit institutions as long as its content is in no way															
CC	modified and this statement is not removed. Usage by and for commercial															
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce															
CC	or send an email to license@isb-sib.ch).															
CC	-----															
DR	EMBL; M24328; AAA29491.1; -															
DR	PIR; A23770; A23770.															
KM	Malaria.															
FT	NON_TER	1	1													
FT	NON_TER	537	537													
SO	SEQUENCE	537	AA	63000	MM	AGD76B15318CC239	CRC64									
Query March																
Best local Similarity	8.7%; Score 597; DB 1; Length 537;															
Matches 212; Conservative 92; Mismatches 193; Indels 246; Gaps 39																
Y	424	NDIDKIKK	KKKGLIN	YKSTLYND	NTINKK	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	DDSSSMV	483								
Db	2	NNNNKNNNN	DDGINYQNT	-NEFKONKK	-NNNPKQYNN	NYKFKEDNNNNNS	NTMSRN	57								
Y	484	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	530								
Db	58	SNVEEH	LRNNSID	MNNNNNNNNNN	NYNQTRFS	FMENENEN	ENKNTYGT	104								
Y	531	DFSSLD	MDMCHLN	NNNDSIFLHK	KLKLDNT	KLKRRSCD	NTMKNKXNLHLARHS	VSGL	590							
Db	105	-----	GNNNN	-----	IHFKKYDN	-----	NNSSKNTDNK	130								
Y	591	TMSYD	POKXND	NTFFKSN	INKMED	NTPKDIL	YESRVN	SMNGNV	645							
Db	131	TDPSTN	-----	MKGTIN	-NDNNNN	MDYLRI	NNININEV	KGSAKKKFTYNN	KNNLKL	179						
Y	646	DISYDEN	NDKIN	NGVILIN	NSNSVS	INSMS	MSINS	MSQ-SNSI	-YKSYNS	MSQS	703					
Db	180	FTONN	NDNNNN	INEDNN	-----	NNNNNNNNNN	NGVFSYQ	NNNNNNNN	NSINIK	MLNNNN	233					
Y	704	ISDQ	QIVVEM	DPNSK	NDNNI	FPDAS	CONNNV	PNITNNNNNNNNNNNNNNNNNNNN	NNNNNNNNNN	NDIVENT	763					
Db	234	INN	-----	NNNNK	SGSQDKNS	-----	NNNFYANN	YVYQ	RKNS	SMNNNNNNNN	274					
Y	764	NNED	GTNS	NKLYA	YNSHNL	FOED	NNKTS	NSINQ	INTNK	KNQDGNV	YNSNFCHY	NLNDK	823			
Db	275	-----	NNNNNN	-----	NHNN	-----	NNNNNN	NHNNNN	NNNNNN	NNNNNN	NNNNNN	NNNNNN	311			
Y	824	NYLID	LNK	KKQK	DKN	IKHG	CONNNI	IQNR	NDE	KKKK	TTYNNNN	VIYNN	NNNN	SGNNNS	PPMK	883

[illegible]

Matches	138;	Conservative	60;	Mismatches	138;	Indels	32;	Gaps	6;
OY	2	NSKRVVRRKRP	SELEKKKDDSIITVYKNNCTIYIDPRKVD-MTKYIERHERIVKV	60					
Db	93	NSNCIVVRKRP	IDDKERQKLDHDSVCFQD-KWHISAKLKVIGITKYLTHNSQDLHT	151					
OY	61	FDDIVDNPFTVENIK	PLIIDLVNENGCVSCFAFGQTSQGTYYMLGSOQPYGQSDTPGIF	120					
Db	152	FGEISTTEQIYLT	PLVPHVSVTQGRATVFCYQIQSSKGTITMNGQ-----	200					
OY	121	QYAAGDIFFTFLNIY	DKNTKGIFFISFEIYCGKLYDLLQRRKMAALENGKEVVYDKL	180					
Db	201	QILAYDIYQGLAEH	TD--LEITVAFFELVSGNVLDDLHGQCRKLTLEDGNGEAVITGLR	258					
OY	181	ILRVLTREELIK	IDVGLRLKIGVNSQDNSSSHALINDLDINKNLSLGAIFDL	240					
Db	259	EVPAITPEPAFQV	IEEGHSILRTQKTEANDSSSHALQVFLRDYGGNLR-GKLGVDL	317					
OY	241	AGSRGADVTYSON	KOTQDTGAINRSLALKECTIRAMDSDKNHIPEFDSSELTVLARDIFV	300					
Db	318	AGSERGSDTKK	NSQKRTESSADINTSLALKECTIRALGQKRAHPYRSKLTLLIKDFS	377					
OY	301	GKSKSIMIANIS	PTISCEQDTLNTLRYSSRYVKNKSTCINEDDTNTERTIILDSKGS	360					
Db	378	PDSKTTWATV	TPSGMSADHSJLNTLRVADRKE-----QVSSNGCRGK	421					
OY	361	EMASSIE	368						
Db	422	AAKASNRE	429						
RESULT 12									
KIP1_YEAST									
ID	KIP1_YEAST	STANDARD;	PRT; 1111 AA.						
AC	P28742;								
DT	01-DEC-1992	(Rel. 24, Created)							
DT	01-DEC-1992	(Rel. 24, Last sequence update)							
DT	15-JUL-1998	(Rel. 36, Last annotation update)							
DE	Kinesin-like protein KIP1.								
GN	KIP1 OR CIN9 OR YBL063W OR YBL0504 OR YBL0521.								
OS	Saccharomyces cerevisiae (Baker's yeast).								
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;								
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.								
OX	NCBI_TextID=4932;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=S288C;								
RX	MEDLINE=92317166; PubMed=1618910;								
RA	Roof D.M., Meluh P.B., Rose M.D.;								
RT	"Kinesin-related proteins required for assembly of the mitotic								
RT	spindle.";								
RL	J. Cell Biol. 118:95-108(1992).								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=S288C;								
RX	MEDLINE=94205666; PubMed=8154187;								
RA	Schertens B., el Bakkoury M., Vlierdeels F., Dubois E., Messenguy F.;								
RT	"Sequencing and functional analysis of a 32,560 bp segment on the								
RT	left arm of yeast chromosome II. Identification of 26 open reading								
RL	frames, including the KIP1 and SEC17 genes.";								
RL	Yeast 9:1355-1371(1993).								
RN	[3]								
RP	CHARACTERIZATION.								
RC	STRAIN=S288C;								
RX	MEDLINE=92354062; PubMed=1643659;								
RA	Saunders W.S., Hoyt M.A.;								
RT	"Kinesin-related proteins required for structural integrity of the								
RT	mitotic spindle.";								
RL	Cell 70:451-458(1992).								
CC	-I- FUNCTION: REQUIRED FOR ASSEMBLY OF THE MITOTIC SPINDLE. INTERACT								
CC	WITH SPINDLE MICROTUBULES TO PRODUCE AN OUTWARDLY DIRECTED								
CC	FORCE ACTING UPON THE POLES. FOLLOWING SPINDLE ASSEMBLY, CIN8 AND								
CC	KIP1 APPARENTLY ACT TO OPPOSE A FORCE THAT DRAWS SEPARATED POLES								


```

CC BACK TOGETHER. THIS FORCE SEEMS TO BE MEDIATE BY KAR3.
CC -1- SUBUNIT: MIGHT BE DIMERIC.
CC -1- SUBCELLULAR LOCATION: SPINDLE MICROTUBULES THAT LIE BETWEEN THE
CC POLES.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BINC
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z13962; CAAB019.1; -
DR EMBL; Z23261; CAAB0785.1; -
DR EMBL; Z35824; CAAB4883.1; -
DR PIR; A42640; A42640.
DR HSSP; P17119; 3KAR.
DR SGD; S0000159; KIP1.
DR GO; GO:0005816; C:spindle pole body; IDA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IPI.
DR GO; GO:0007020; F:mitotic anaphase B; IMP.
DR GO; GO:0000092; P:mitotic anaphase B; IMP.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin_1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
DR Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
KM Mitosis; Cell cycle.
FT DOMAIN 50 423 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 424 510 COILED COIL (POTENTIAL).
FT DOMAIN 648 670 COILED COIL (POTENTIAL).
FT DOMAIN 710 780 COILED COIL (POTENTIAL).
FT DOMAIN 808 828 COILED COIL (POTENTIAL).
FT NP BIND 141 148 ATP (BY SIMILARITY).
SQ SEQUENCE 1111 AA; 125794 MW; 212F8279766137FC CRC64;

Query Match 8.2%; Score 561.5; DB 1; Length 1111;
Best Local Similarity 22.9%; Pred. No. 1e-17;
Matches 255; Conservative 204; Mismatches 393; Indels 263; Gaps 48;

QY 2 NSKIKVYVRKRPSELEKKKSDITV---KNNCTIYIDPRKYKDMTYRIERHYVD 58
DB 50 DSNIHVYVRCRSNKRKEIEKSSVISTLGPQKEIILNSGSHQSYSSK-KTYQF--D 105
QY 59 KVFDDTVDFVNFENTIKPLIIDLYENGCVSCFAYGQTSKTYTMG-----S 108
DB 106 QVFGBSDQETVFNATKNIKEML-HGVNCTIFAYQGTGTGTYTMSGGINLGDVQST 164
QY 109 QPYGQSTPGIFQYAGDITFTLNYDKNTKGIFFISFYEGKLYDIQ----- 159
DB 165 DNLIGHHAIIIPVLDLFEKLSLNKEYS--VKISFELVYENMLKDLSDSEDDPAV 222
QY 160 ---KRRK-VAALNGKKEVVVKILIRVLTKRELKIMLDVYLKIKIGVNSQNDSSRS 215
DB 223 NDKRQRIPIFDNNNNNSINWKNQELFINSAHEGLMLQSGSLKRRVAATKCNDSRS 282
QY 216 HALINTI-----DLKD--INKN-TSLGKIAFIDLAGESEGAPTVSNKOTOTDGANIN 265
DB 283 HTVFTITTNVEEDSDKHGQNKFPVKIGKLVLDIAGSEINNSGAEKRAQAGL-INK 341
QY 266 SLALKECTIRAMDSKKNHPRPDSLTQVLDIFVGSKSIIMANISPTISCEQTLNTL 325
DB 342 SLTLTGAVINALVDHSHNPIYRESKLTRLQDSLGWTCTCIATISPAKISMEETASTL 401
QY 326 RYSSRYVKNFKN-----KSCVINE-----EDDNTERTISLDSKSEMA 364
DB 402 EYATRAKSIKNTPOVNNOSLSKDTCLADYIOEIKRLNDLKNRKNKGIFITQ---DQLD 458

```

```

QY 365 SSIEYVVKSNHLLSNNNNNKINRGKIN--DKIERNNILKN-----KSPDKREGEF 413
DB 459 YESNLSLIDBQNKILNLEBQIKKFKENTANQDNLNLLQSEKXILAIQNNVDFSNF 518
QY 414 TSTFGK--YSLNDIDIKKKKKGLINRYKSTLYNDN-----TINKKNNNN 458
DB 519 YSEIQIKHHTNLEINNEVILQORDPSLENSQK-QYNTNQMWQLKISQOVLQTLNTLQSLN 577
QY 459 NNNNDND-----NNNDNNNNNNNS--SMVNNMIMHNNNNNNNI--NVNNNNNN 507
DB 578 NYNSKSEVYIKGTEELTRVNTTHKAKHSTLSKLSLINTTLNMQNLEIVSISTSL 637
QY 508 NNNNNNSH--NHLPOPNYAF-----TTPSPSSLDWM--NCHLNNDK 547
DB 638 FQSDTSHYRKDLNLEYQHQFLKVLQNDIKSCDLSISSILTSINELISQNTTLNSM 697
QY 548 STFLKKNLDNLIKLNKSSCDNINAKKNNL-----HLAR- 583
DB 698 NVLIENQSGSSQLIEQ--DLEIKKLNDLINERRISNQFNQQLAEMKRYFQDHVSRT 754
QY 584 ---HSVSGKLTWFSYDPQKNKNTFFKSN-----INKGENDTPKDIYESRN-- 627
DB 755 RSEFHDELKNCIDNLKQKQSKLDQDIWQKASIFNETDIYVNRKHSDSIASLAHMENTL 814
QY 628 --VSNNK---GNVLGLANKYTHHDISTK---DENHNDKINNGVINIINNSVNSINNS 678
DB 815 KTVSQNNBFTDNLISLSGMMNDISSKLRSLPINFLKISGTCICETGDDTTA-SNP 873
QY 679 NNNSINN-SMANSNSIYKSNYNSNOSISDVQIRYVNEMDT-SKNNNDNIFPDALSCDNN 736
DB 874 VLTSLKPPONILICSDIALTNEKIMSLIDEIQ---SQIETISENNINL----- 918
QY 737 YPNTNNNNNNNNNN-----NNNNNIDVENYNNNRDGTNNSMKLYA-----YNSHNL 784
DB 919 ---IATNEPNSICNFLTIDYBENTWQISKTDDEVLSCECKLQSKILGMDIFTHASLE 975
QY 785 QP---DNKNKTSNIQNTNKNQODGNVYSMNFCHYNLNDKRYLLIDLNNKEQK-KNH 840
DB 976 KPLHEHTREPAVITKALPL-----LDYPQGFQYR-----DAKNSKDDTSNR 1019
QY 841 GCDNNITIQNRN---DPEKK-----KKTNFPNNNNNIVYVNNMG----- 875
DB 1020 TCIPLNSTENPPLSFQSPKTPVPVDPQPLPKVLLPKSINSAKNSKSLTPYTEGGRS 1079
QY 876 NNSPKMYKGLGSHSISDMKNEMKNEMDNE 910
DB 1080 QNNLKR-----RFTTEPLKGEETENNDDIQNK 1107

RESULT 13
RPCL_PLAFA
ID_RPCL_PLAFA STANDARD; PRT; 2339 AA.
AC F27625;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA-directed RNA polymerase III largest subunit (EC 2.7.7.6).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Hemosporida; Plasmodium.
ON NCBI_Taxid=5833;
RX SEQUENCE FROM N.A.
RP MEDLINE=92018020; PubMed=1656254;
RA Li W.B., Baik D.J., Tanaka M., Gu H., Fox B.A., Inselburg J.;
RT "Characterization of the gene encoding the largest subunit of
RT Plasmodium falciparum RNA polymerase III."
RL Mol. Biochem. Parasitol. 46:229-240 (1991).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -1- SUBUNIT: RNA POLYMERASE III CONSISTS OF ABOUT 15 DIFFERENT

```



```

Db      1227 HRTNDI-QPDLCOILNKDPRNVAVLTGKAFIFLKKKPYSPHLPYEECKNIVHYIMK 1285
      : : : : | : : : | : : : |
Qy      803 -----NNOGANNVYSNMFCGYINLNDKRYLIDLNNKCK-DKIHICGDNNTION- 849
      : : : : | : : : | : : : |
Db      1286 KGHKKIKININNNHNSNLYY-----HYNIIDT-FVKRNKNGMCKNLLYKIQOQLVNL 1338
      : : : : | : : : | : : : |
Qy      850 -RNDPEKK-KTNFNNNNIYIVNNNMNNNSPRMKYGLCSHNSIDMKNNKMMKNEMK 907
      : : : : | : : : | : : : |
Db      1339 IHNLYKKKKYNNNYIDIDEVHLGNNNNNNN-----KNNSEKKPL 1380
      : : : : | : : : | : : : |
Qy      908 DNEKMDNHIKSN--NNNSSSSSSSSNNNI-----YNNINDDTFONDYCHNDYFTIRKKN 961
      : : : : | : : : | : : : |
Db      1381 KKKCK-HIRKNESNDITFTNTYSNNIHLKSKYKVHHKNYYPDSCIN-----LRKKCN 1433
      : : : : | : : : | : : : |
Qy      962 T--NINSNTYQNDITITNSLNDYKSNITLHFEKKTYYFTLSTNEDI-YNKEMEGKI 1017
      : : : : | : : : | : : : |
Db      1434 SLFYNLKKYIYER-----KKYLOHCLLK-HDNKYKVELPRIXDINYSQWESIKT 1483
      : : : : | : : : | : : : |
Qy      1018 R-----LDDODKYD-----NDNNVNNN-----NKNVNDNNVNDNNVNDNDX 1059
      : : : : | : : : | : : : |
Db      1484 RNFHHSLEQFAFNSLILSPYIKDDNNVYKNTYKNTYIN--KNSICNRYICNK 1540
      : : : : | : : : | : : : |
Qy      1060 NNVNNVNDNDDDVDVFNHINKNFNNNEYLASYOKNVDPTINNCLNSLDISMYDTKEIL 1119
      : : : : | : : : | : : : |
Db      1541 NYIYNKN-----NIVYKN-----INKNKILTHAKSVL-----LSGSSKKELKF 1581
      : : : : | : : : | : : : |
Qy      1120 NNILSKYKAEKDNVYKTYINEDIKMSLEIIDKTAQSIYEKKYLLTKLILFKQNDV 1179
      : : : : | : : : | : : : |
Db      1582 SNIRRHKLKEKKN-----KKNIKRYKNNVNNNTS-----KGHILMCTGHEFKDYSS 1630
      : : : : | : : : | : : : |
Qy      1180 Q-----INNETSDLRKDLVWCHICNNNPDQPHYAVSRLEKDIINLIMLRQIWCESN 1233
      : : : : | : : : | : : : |
Db      1631 LKAKRYIYNNKRYMLKNDV-----YDRMNMILTMYRGTQIGCSK- 1672
      : : : : | : : : | : : : |
Qy      1234 LRLLYQLFVVEYONKSANSVILLNVSNNGDIILNK-----KIVQDNIKNSMDH 1282
      : : : : | : : : | : : : |
Db      1673 -----KVK-----NIYNNNNNNILKKIKIRPLEHLLVDCKKNICH 1708
      : : : : | : : : | : : : |

RESULT 15
CUT7_SCHPO STANDARD; PRT; 1085 AA.
ID_CUT7_SCHPO
AC P24339;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Kinesin-like protein cut7.
GN CUT7 OR SPAC25G10.07C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX Hagan I., Yanagida M.;
RT "Novel potential mitotic motor protein encoded by the fission yeast
  cut7+ gene."
RL Nature 347:563-566(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; Pubmed=11859360;
  Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
  RA Sgouros J., Peat N., Haylee J., Baker S., Basham D., Bowman S.,
  RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
  RA Gentles M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
  RA Gellies S., Goble A., Hamlin N., Harris P., Hidalgo J., Hodgson G.,
  RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
  RA James K., Jones L., Jones M., Leather S., McDonald S., Mclean J.,
  RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

```

```

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbinoletsch E.,
RA Sherrford K., Rutter S., Sanders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voiclaert G., Aert R., Robben J., Grymowicz B.,
RA Welfans J., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs K., Firtz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leinach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Mamut R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelare V., Mortier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daea R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cernuti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Usery D., Barrett B.G., Nure P.,
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -1- FUNCTION: COULD BE A SPINDLE POLE BODY MOTOR. ON TRANSITION FROM
  G2 TO M PHASE OF THE CELL CYCLE, THE SPINDLE POLE BODY DUPLICATES;
  THE DAUGHTER POLE BODIES SEED MICROTUBULES WHICH INTERDIGITATE TO
  FORM A SHORT SPINDLE THAT ELONGATES TOWARD THE NUCLEUS AT
  METAPHASE. MUTATIONS AT CUT7 BLOCK SPINDLE FORMATION.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BINC
  SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.ebi.ac.uk/announcements
  or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; X57513; CA440738.1; -
DR EMBL; Z70691; CA94636.1; -
DR PIR; T38378; T38378.
DR HSSP; P17119; 3KAR.
DR GenedB; Spombe; SPAC25G10.07c; -.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin_1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; Kisc; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
KW Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
KW Mitosis; Cell cycle; Phosphorylation; Repeat.
FT DOMAIN 70 435 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 436 604 COILED COIL (POTENTIAL).
FT DOMAIN 715 740 COILED COIL (POTENTIAL).
FT DOMAIN 897 955 COILED COIL (POTENTIAL).
FT NP_BIND 159 166 ATP (BY SIMILARITY).
FT REPEAT 987 998
FT REPEAT 999 1010
FT MOD_RES 1011 1011
FT FT
FT COMPACT 34 61 PHOSPHORYLATION (BY CDC2) (BY
  SIMILARITY).
FT FT SASNPKRREPITDYGPRSDNSPT ->
  LRALINDVSLILLT (IN REF. 1).
SQ SEQUENCE 1085 AA; 122133 MM; 5669277875559D58 CRC64;

Query Match 7.9%; Score 538.5; DB 1; Length 1085;
Best Local Similarity 20.7%; Pred. No. 1e-16;
Matches 235; Conservative 223; Mismatches 410; Indels 265; Gaps 43;

Qy 3 SKIVYVRKPLSLKXKSDIITYKNNCTLYID-----EPKRVDMTKIERHEP 55
      : : : : | : : : | : : : |
Db 71 TINVAVVRKRTIOE-----VRDNSLAVSTSGMGALALAQSPSSMLVTKTY 120
      : : : : | : : : | : : : |
Qy 56 IVDVFPDVTNFTVYNTIKPLIDLYENGCVSCFAYGOTGSGKTYTMGSPYQSGD 115
      : : : : | : : : | : : : |
Db 121 AFDVYFGEADQLMFLFNSVAPMLEQVL-NQYNTCTIAYGGTGKTYTNSGD---LSD 175
      : : : : | : : : | : : : |
Qy 116 TPGIFYAAG---DIFFLNIYDKNTK-GIFISFYEYCGKLYDL---QKRNVAAAL 167
      : : : : | : : : | : : : |

```



```

Db 176 SDILSEAGLIPRALQULFSSLDNSQGEYAVKSYELEYNEIRDLVSEELRKPARYP 235
168 EN-----GKEVYVAD-----LKLIVLTKMELILMIGVILRKIGVNSONDESS 213
236 EDTSRGVAVITIGIEEYIKNAGDGLRLR-----BSHRQVAAKCKNDLSS 283
214 RSHAILNIDUK-----DINKTS-----LGKIAFDLAGSERGADTVSONKO 255
284 RSHSIFITILHRAVSSGMDTETNSLTINNSSDULRASKLHMVDLAGSEMGSGAENKR 343
QY 256 TQTDGANIRSLALKECIPAMSDKNHIFPDSLTGLRDI FVGKSKS IMANTISPT 315
344 ARETGM-INOSLTLTGKIVINALEKHAHPIYRESKTLRLQDLSLGGTKTSMITVTVSSTV 402
Db 316 SCEQTLNTLRSSRVNFKNGSTCINEEDTTERISI-----LDSKSENNASSIEN 369
403 TNEETITISLEIYARAKSIRNKP-----QNNQVFRKVLKOLVDLIRLKNDLNATRKKN 458
QY 370 VV-----IKSNHLGNNNNNKINRGINDKIERNNI--LKNKSPDKPREGPT 414
459 GYLAESTYKELMDRVQONKDLCOEQBARKEVDLNVKSSRELOQYVSKSNGEKKVEVA 518
Db 415 STFGKTSILNDIKIKKAKKGLINTYSTLYNDNTINKKNNNNNNNN-----462
519 LQLOLVNSSSTELSVKSENEK-----LKNELVLEIEKKKETEAKITTVATDLSQYRE 574
QY 463 -----DNNNNNNNNNNNNNNSSVMNNMINH--INNINNINN--VNNNN 506
575 SKEIYASLYEKDRTERRNNKENNNWNLTLMRSPHSGFTBETNGYFTLLDFNA 634
QY 507 NNNNNNNNNHNLPOPNVAFD--TSDFSLDD-----MCHLANNNDKSIFLHKNLARN 559
635 SWEELANTHSNOL--LISWTKITEHQSDALQASRSCAIPNS--SLDLVSLKXO--688
QY 560 IKAKNRSSCDNIMKKKNNIHLARHVSGLTMTFSYDPOKKNKNTFPKSNINMEEDNTPK 619
689 ---SKNSLIDALEHSLDIDISMSQKLGNGISSELELOKDMESY-----R 731
QY 620 DILYESRNVNNGNVLGLNKRTHTDISTRKNNHNDKINNVININNSNVNSINNSN 679
732 QVQELRSLNLYL-----QHTHEE-----SOKELMYGVNR--DIDALVKTC 769
QY 680 MNSINNSNM--NSNSIYKSNVNSNO-----SISDQIRVYEMDTSNKNNDIFPDALS 731
770 TSLNADIIISLYIDQSKFESKQODLIANGKIVSNFLOE-----QNESLYTKA--821
QY 732 CDNNMTPNITNNNNNNNNNNNNNNIIVENYNNRDTGNNSMKLVAYNSHNLFQDPNNKN 791
822 -----DILHSLNDTNSIRKANEI-----MNNRSEFLRNA-----853
QY 792 TSNIGNINNNKNNQDGNVYSMNFCHYNNLNDKRYLLDANKKQKQKIHGCC-----NN 845
854 ASQALIVGANKEKRIQTVENGSQL-----LDSKSKALHSNRSRYDHCLALAESQKQGVNL 909
QY 846 IICNRNDFEKKKTNFYNNNN-----IYVNNNMGNNSNPRMYKGLCSHTSIDMKKN 899
910 EYQUTLRLQKYVEHEDMTKEGQQLDLLESLVGNND-----LIDISIKTP 957
Db 900 EMKNNEMKONEMKONHKSNNNS--SSSSSSNNNINNNINDDTFONDYCHANDTFTR 957
958 HTELQKITHLVLTGTSILANHTHELGLDESLCNLETTIEDTSLVK-----LETTGDP 1012
QY 958 RKNNTNINSNIYONDDIITYINSIYNGMNTLHFXEKTYPPLSTNEDINYK 1010
1013 SKRELPAFSPWTRDSSLIKETTLN--LSDKKFVREYITSSNQINEPVIDK 1063

```

RESULT 16
K125 ARATH
ID K125 ARATH STANDARD; PRT; 1056 AA.
AC P82266;
DT 16-OCT-2001 (Rel. 40, Created)

```

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DB Probable 125 kDa kinesin-related protein.
GN ATG36200 OR P2H17.19
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia; PubMed=10611797;
RX MEDLINE=20083487;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Bento M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H.L.,
RA Moffit K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Unayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.V., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana."
RL Nature 402:761-768 (1999).
CC -1- FUNCTION: RESPONSIBLE FOR MICROTUBULE TRANSLLOCATION. MAY BE
CC IMPORTANT FOR THE ORGANIZATION OF PHRAGMOPLAST-SPECIFIC ARRAYS OF
CC MICROTUBULES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIOC
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, AC006921; AAD21445.1; -.
DR PIR; H84777; H84777.
DR HSSP; P17119; 3KAR.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
KW Motor protein; Microtubule; ATP-binding; Coiled coil.
FT DOMAIN 1 363 KINESIN-MOTOR (BY SIMILARITY).
FT 406 526 COILED COIL (POTENTIAL).
FT NP_BIND 98 105 ATP (BY SIMILARITY).
SQ SEQUENCE 1056 AA; 119267 MW; 7AD18A3FA9A9211 CRC64;
Query Match 7.6%; Score 520.5; DB 1; Length 1056;
Best Local Similarity 21.6%; Pred. No. 6-16;
Matches 240; Conservative 207; Mismatches 424; Indels 241; Gaps 42;
QY 5 IVVVRKAPLSELEKKKDDSIITYKNNCTLYIDSPRYKVDNT-KYIERHEFTVDKVPDD 63
DB 13 VQVLLRCRPFSDDELRSNAPQVLT-----CNDLQREVAVASONAGKHIDR-VTFPKVGP 67
QY 64 TYDNFTYVENTIKPILIDLYENGCCSCGFAYGQGTSGTYTMLG-----SQPYQSDPT- 116
DB 68 SAQQKDLIDQAVAPLVNEVL-GENCTLFAVGQGTGTGTYTMEGCRSRKSPKCGGLPAE 126
QY 117 PGIFQYAGDIFTFLINVDKNTKGIFFSYEYIGKLYDL-----QKRKYV 164
DB 127 AGVIRPAVKQIFDLE--GQAQVSVKTVLELVNHEITDILAPEDLSRVAAEKKKPL 184
QY 165 AALENGKEVVVKDKIKIRVLTKEELILKMTDGLRKIGVNSONDESSRSHAILNIDUK 224
DB 185 PLMEDGKGVAVRGIEEIVTSANEIPTLLERGSSSKRTAETFLNKQSSRSLSFSIT- 243

```



```

QY 225 DINKNT-----SLGKIAPIDLAGSERGADTVSOKOTQTDGANTIRSLALKECIRAM 277
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 244 HIKATEPEGEBELIKGKLTINLVDLASGENISRGDRARERAG-EIKSLITLGRVLSAL 302
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 278 DSDKNHLPFRDSELTKTLARDIFVSKSISIMANISPTISCCEOTLNTFRYSRVKFNK 337
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 303 VEHGHVPIRDSKTLRLRDSLGRTYCTIATVSPAVHCEETLTLDLVAHRKAKIRNK 362
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 338 STCINEEDTNT-----ERISILDSKSENNASIIEN-VVIKSNHLLISNNNNKIN 387
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 363 PE-VNQMKESTLIKDYGEIERLK-----AEVYASREKGVMPKERYQEESERKVM 415
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 388 R-----GKINDIEKNNIANKSPDKPRE-----GFTSREKYSIADLID 427
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 416 AEOIEQWGGQIENQOKLELODKYVGQRECSDLTKLIDITEKNLQTCVKLASTN--E 473
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 428 KIKKN-----KKKGLI-----NYKSTLYNDNTINKKNNNNNNND 463
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 474 ELKKSQYAMKEKDPITISQKSENVLVQOACILQSNLEKATKDNSLHOKIGREDKXSAD 533
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 464 NNDNNNNNNNNNNNDSSVMNNMINNNNNNNNNNNNNNNNNNNNNNNNNHNLPOPN 523
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 534 ---NRKVVNYQYELSEQISNLFRRVASCLOQNVHLQGVNKLQSRLFAHKKALIEMK 589
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 524 YAFDTSDPSLDMNCHLNNDKSIPLHKK---LRDNIKLKKRSSC-----DN 570
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 590 KKVASRDLYS-----SHLEAVQVNVRLHKKANANACEEVALTTSSACSIDEFLASGE 644
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 571 INMKKNNHLARBSVGSKLTIMFSYDPOKNKDNTPFKSINKMEDNTP---KOLYESRN 627
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 645 TTSSLFDELQSLASHQGEMLFARBLRQRPHTT--EQTEMBERYSTPFQKLMESKN 702
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 628 VSNMNGVVLGLAKNKTTHHDISTKDNHNDKINNGVINI INSNVNSINNSNNNSINN 687
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 703 A-----ETRAEANDSOINS-IIDQKTYEAKOSKSDTKLADLTN 742
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 688 MNSSIYSKSNVNSNOSISDVQIRYVNEEDTSNKNDNFFPAILGCDNNMYNINNNNNN 747
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 743 LVSS-----HIRQHELVDLSRLNH--FKQAVSSNKTFELDEHVASVNNL 783
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 748 NNNNNNNNNNNIDVENYNN-RDGTNNSMGLYANSHNLFQDPNNKNTSNIQNTNKNQD 806
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 784 TKDAKRRKETSQMOENAREGADSSAKHCMEILLQOSVGHASA-----830
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 807 GNVVSNMFCYNNLNDKXYLIDLNNKEQDKNI---HSCDNNIIONRNDP-----KKK 857
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 831 -----FKKCKITHBS-LKEMTSKQVTVSVLSVACSNS---EQHDAEVDARSATAE 878
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 858 KTNFYNNNNIYIVANNMGNNSNPPRKVGLCSHTSIDMMKNEMKN-NEMKDNEMKXHI 916
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 879 KDVTNKSDDII-----QOIERMSDEBKASVSKEIEN-----V 910
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 917 KSNNNSSSSSSSSNNNNIYNNIND--DTPONDYCHNDTFTLIRKQNTNINSNIYQNDI 974
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 911 RSHKELTSPQDDCCQARCIEDKAOETFOOQYMEYEPFGALPTTGNBEIPIPKA-----964
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 975 IYTTNSLNDYMSNTLL-HPEKITYPTLSTNE 1005
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 965 --TISLRAMPJETLVEEFRENNSYESPATKE 994
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

OC OC Spermatothya; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_Taxid=4097;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 149-163 AND 194-211, AND
RP CHARACTERIZATION.
RC STRAIN=cv, Bright Yellow 2;
RX MEDLINE=97196959; Pubmed=9044048;
RA Aada T., Kuriyama R., Shibaoka H.;
RT "TRP125, a kinesin-related protein involved in the centrosome-
RT independent organization of the cytoskeletal apparatus in tobacco BY-2
RT cells."
RL J. Cell Sci. 110:179-189(1997).
RN [2]
RP FUNCTION.
RC STRAIN=cv, Bright Yellow 2;
RX MEDLINE=95074298; Pubmed=7983184;
RA Aada T., Shibaoka H.;
RT "Isolation of polypeptides with microtubule-translocating activity
RT from phragmoplasts of tobacco BY-2 cells."
RL J. Cell Sci. 107:2249-2257(1994).
CC -1- FUNCTION: RESPONSIBLE FOR MICROTUBULE TRANSLLOCATION. MAY BE
CC IMPORTANT FOR THE ORGANIZATION OF PHRAGMOPLAST-SPECIFIC ARRAYS OF
CC MICROTUBULES.
CC -1- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING S PHASE. EXPRESSION
CC INCREASES AS CELL MOVES FROM S PHASE TO M PHASE AND THEN DECREASES
CC RAPIDLY AS CELL ENTERS THE G1 PHASE. EXPRESSION INCREASES AGAIN
CC DURING THE G2 PHASE.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. B1MC
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
DR EMBL; D83711; BAA23159.1; -.
DR PIR; T02017; T02017.
DR HSSP; P17119; 3KAR.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; Kisc; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
KW KINESIN-MOTOR.
FT DOMAIN 1 359
FT DOMAIN 2 371 522
FT NP_BIND 95 102
FT COILED COIL (POTENTIAL).
FT NP_BIND 150 150
FT T->Q (IN REF. 1; AA SEQUENCE).
FT CONFLICT 153 153 E->F (IN REF. 1; AA SEQUENCE).
FT CONFLICT 161 161 D->Q (IN REF. 1; AA SEQUENCE).
SQ SEQUENCE 1006 AA; 113710 MW; 890C0E0F3504AA7D CRC64;

Query Match 7.5%; Score 513; DB 1; Length 1006;
Best local similarity 20.2%; Pred. No. 1.2e-15;
Matches 234; Conservative 209; Mismatches 379; Indels 338; Gaps 46;

```


QY	165	AALNEGKXEVVVLDKILRLVLTJEEELMKMDIGVLRLKGVNSQDSESSHALINTDLK	22
Db	181	PLMEDGKGVVLVGBEEELVTSANEFLTLEEGSAKRRAETLTKOSSRSHSLPSTI-	23
QY	225	DINKNT-----SLGKAFIDLAGERGADTVSQNKQOTDGANINSILALKECIRAM	277
Db	240	HKEATEGEBEELJKCGKTLNVLDSAGEENISRGAREGRAREAG-EINRSILLTKGVINAL	296
QY	278	DSDKNIHPPFDSELTXYLADIYFGKSKSMIANISPTLSCCQTLNTRPYSRYKFNFK	337
Db	299	VEHLGHIPYRDSKLTIRLSDLSGRTKTCIATVSPAHCUEETLSTIDYAHRAKNIKX	356
QY	338	STCINEEDDINTERISILDSKSEMMASSIEMVVIKSNHLJSMNNNNKNIKNGKINDKIER	397
Db	359	P-----EYQKQMK	367
QY	398	NNILKNKSPDKPREGFTSPFGKXSLNDIDIK-----KNKKGLINYKSTLYNDNTINK	452
Db	368	STLKD-----LYG-----EIERKAEVYAREKGVYIPEKRYGE-----	408
QY	453	KHNNNNNNNDNDNDNNNNNNNDSSSMVNMNIHNNNNNNNNNNNNNNNNNNNNNN	512
Db	405	-----ENERKMA-DOIEQM-----GVSIEHQHOFELQ	433
QY	513	NSHNHLPPENYVAFPTDPS-SLYDMNGHLLNNNKSIFFLHKNLRLD-ILKXRSSCDN	570
Db	434	SRHDSQVQ-----CSDLTCLDVTQQLNQTSLKLAAYTEBQLRQGYTLTKERDPLIS	486
QY	571	IMNKKKNL-HLA-----RHSVSGKLTMSYDQKKNKDTFFPKSNKKMEDTTPKDIL	622
Db	487	EOKKAEVALAHQCVLRADLEKSIQENASLFOKAREDKLSTDNRSILVNFOELAKOL-	545
QY	623	YESRNVSNNGNVLLGKNTKTHDIDSTCKBNND-NKINGYINININNSVNSINNSNM	681
Db	546	-----GSLESTLATSVCQTEHLQCEVFXCNFLDSHDKALDLKRIKIN-----	589
QY	682	SINNSNNNSNIYKSNVNSNOSI-----SDVOIRYVENMDTSKNDIIFPALISD-	733
Db	590	-----SSMALYISHFEMQNVYLRKHTSNATLEBVSCLASSNISITKEFLDAEAVEA	642
QY	734	NNMYPNITNNNNNN-----NNNNNNNNIID-VENYNNRGGTNNSMKLYA	777
Db	643	NSMFELOSTLSTHQEMAHFARELRQRFENDSTELHTNISALIQRFDPK-LLDESURL-	699
QY	778	YNSHNLFPQPNKKNTSINQININPK--ANNQDGNVSNMFCGYNLNDGYLIDLANKQK	835
Db	700	-----EKHATTVEIQTNSIAEFKAYEEOQSDEKLLADVTSLVSNHMRQK	748
QY	836	D-----KNHGCNNIILONRNP-----EKKKKNF-----	861
Db	749	EIVGARLVDLRETVSGNRTFLDGHVSSMGITTDKAKKQDPFYMAQBEETEFENADPFAK	808
QY	862	-----YNNNNIYVNNNGNNNSPMKYGKCGSHTSIDNN-KN	898
Db	809	HCEMESIMQCVSTAETALRKQSTHELV--NDMGNOHVLTM-----HSVRLICDN	858
QY	899	NEM-----KNMKNKNEKDNH--IKSNNNNS-----SSSSNNNIYNNIND	939
Db	859	NEQHVTFDSTRESAEDVKNRSEDITKSIDLSGEBRGSJSGVLDTSSAHSFTLLVLKK	918
QY	940	D-----DFFONDYCHNDTFFIRKKNNTINSNIYQNDIITYINSLANDYMS	966
Db	919	DHCQGSTSIQIALTEFFQOKYMDYETPTATPIRSEDPVSKY-----TIEBLRAMPM	970
QY	987	NTLL-HPEKXYTPTISTNE	1005
Db	971	EVLLEERENNSSESFOVKE	990


```

Db 170 ERFVSSPEEVEVIDEKGKNNHVAVTNNHNSRSHSVFLINVKQENVEYQKLSGLYL 229
QY 238 IDLAGSERGADTVSONKQOTQDGNINRSLALNECTRAM-DSDKNHPRPDSLTKLR 296
Db 230 VDLAGESEVSKTGAEG-AVLDEAKINRSLALGNVSLADGNSHVYRDSKLTLLQ 288
QY 297 DIFGKSKSIANISPTISCECEQTLNTRYSRRVKNKSTCINEB-----DDT 347
Db 289 ESLGNARTMTVICSPASVNESKSLFLGQAKTIKQ-VVSNBESLTDDEMRREK 347
QY 348 NTERISILDSKGESEMASSIEENVIKSNHLLSNNNKINRGKINDKI--ERNNILNK 404
Db 348 EKEVTKLKATMAKLEA-----ELQWRKTCQAASVEEYQDLKEDV 387
QY 405 SFDPREBGFST-TRGKYSLLNDIKIKKKKKGLINYSTLYNNTINKK----- 453
Db 388 PAESPATYSTLAGGLASMEGDTQLEER-LKYQDLDDKDEINNOQLIEKLEQ 446
QY 454 -----HNNNNNNNNNDNNNDNNNNNDSSVNNMIMNNNNNNNNNNNN 504
Db 447 MMEQEDLIAOSRDEYENIQODMSRQADNESAKDEVEVLOAEELANVYQKGEVDEK 506
QY 505 NNNNNNNNNNNHNNHLPQPNVYAFPTDTSFSLDDMNCILNNDKSTFLHKKRLDNI--KL 562
Db 507 NKEHENLSEELNOKL-----STINSIQNEILDQKDSMHRKKEVTDMMINIL 553
QY 563 KNRSSCDINMKKKNNLHLARHSVSKLTMFSYDPQKKNDTFPKSNINNKEDNPKOIL 622
Db 554 KDLADIGTIVG--GNALETPTAGS-----GEKIEEFTVALYISKMS----- 596
QY 623 YESHVSNNNGNVLLGLNKNTHTDSTDENHNDKINNGYININNSVNSINNSNNS 682
Db 597 -EVTIVSRNNQL-----ENTQODNFKIETHKXOLSKCKLLOQHEKMASSLOALIDS 650
QY 683 INNSNNSNSTYKSNYSNOSISDVQIRYV---NEWDTSKNDNIFPDALSCDNNYP 738
Db 651 ENKRMLEDNVDLSIEEYAKLKAQOMHIALSEREKTSQASETR--EVLKQMEMHR 707
QY 739 NITNNNNNNNNNNNNNNNNNNNNNNNNNDGTTNSM--KLYAYNNHNLFGQDNNKNTSNI 795
Db 708 E-QHQKQLOSLRDEISEKQATVDNKD-DNORLSIALBKLAQ--DYDKLKQEEVAKAKL 763
QY 796 QNITNKNKQO-----DGNVYSNMFCHYNLDKNVLLDLNNK-----EOKDK 837
Db 764 ADLSGIQDRRQAKODLGLKETVAKELQTLH-NLR-KLFPYODLONKXKKSCKTEEBDE 821
QY 838 NIHGCDNNIIONRDNFEKKKKTFFYNN--NIVIVNNNGNNNS-----PRMKYGLCS 889
Db 822 DTGG-----NAAQOKISFLENNLEQLTKVHKLQVLDNADLRCELPLERKL--- 868
QY 890 HTSIDNNKNNEMKNNEMKDNEMKD 913
Db 869 RATERVRSLSALKAKEGANRD 892

```

```

RESULT 19
BIMC_EMENT
ID_BIMC_EMENT STANDARD; PRT: 1184 AA.
AC P17120;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE kinesin-like protein BIMC.
GN BIMC.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=162425;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9019985; PubMed=2138511;
RA Enos A.P., Morris N.R.;
RT "Mutation of a gene that encodes a kinesin-like protein blocks

```

```

RT nuclear division in A. nidulans.";
RL Cell 60:1019-1027(1990).
CC -!- FUNCTION: IMPORTANT ROLE IN MITOTIC DIVIDING CELLS. MICROTUBULE
CC MOTOR REQUIRED FOR SPINDLE BODY SEPARATION.
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIMC
CC SOBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32075; AAA33298.1; -.
DR PIR; A34795; A34795.
DR HSP; P17119; 3KAR.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin_1;
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN; 1.
KW Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
KW Mitosis; Cell cycle; Phosphorylation.
FT DOMAIN 79 485 KINESIN-MOTOR (BY SIMILARITY).
FT 489 900 COILED COIL (POTENTIAL).
FT DONAIN 901 1184 GLOBULAR (POTENTIAL).
FT NP_BIND 167 174 ATP (BY SIMILARITY).
FT MOD_RSS 1006 1006 PHOSPHORYLATION (BY CDC2) (BY
FT FT SIMILARITY).
SQ
SEQUENCE 1184 AA; 131630 MW; CEAD1SEC8F980E4F CRC64;

Query Match 7.3%; Score 500; DB 1; Length 1184;
Best Local Similarity 21.1%; Pred. No. 5.3e-15;
Matches 253; Conservative 194; Mismatches 415; Indels 338; Gaps 48;

2 NSKTKVVRKRPSELEKKKSDIITVKNNCTLYDEPRYKVM-TKYIERHEFYDVY 60
Db 79 DTSIHVVVRCHGRNEREYKENVGLQTEG-----VAGKYVELSMGRNANVKTYTDDKV 133
QY 61 FDDTVDFVTEVNTIKPLIIDLYENGCVCFAVGQSGKTYTMLGSOYPGQSDTPGIF 120
Db 134 FSAADQITVEYEDVVLIVEMLA-GYNCITFAVGQGTGKTYTMSGD---MDTILGIL 188
QY 121 QYAAGD-----FTFLAIYDKONTKGIPISTYEYCKGLDILL-----OKR 161
Db 189 SDNAGIIPRYLSLFAKLADTEST--VKCSFTELYNBELRDLAENPNKLIYDNEOKK 246
QY 162 KVAVALENGKKEVVVKD-----LKLIRVLTKELILKMIIDGVLRKIGVNSQNDSESRSHA 217
Db 247 GHWSTLVQMEETTYIDSNATGIKLQ-----QSHKQVATATKNDLSSSH 294
QY 218 ILNIDLDINKNTSL-----GKIATFDLAGSERGADTVSONKQOTDGNINRSLAL 270
Db 295 VFTTV-NIKRTTSGEYVPGKLNVLDAAGSENGRGAENRA-TEAGLINKSILLTL 352
QY 271 KECIRPANDSKNHIPPRDSLTKVLRDIFPGKSKSIANISPTISCECEQTLNTRYSR 330
Db 353 GRVINALVDSQHPIPYRESKLTLLQSLGRTTCIATMSPRSMILEFTISLDAFR 412
QY 331 VKNFKNSTCIN-----BEDPTNERISILDSKGESEMASSIEENVI--KS 374
Db 413 AKNIRNKPQ-INSTMPKQTLIREFTAEIKLKAELIATRRHNGYVMSVESTEEMKNEES 471
QY 375 NHTLSNNNNKIN--RGKINDKIERNNILKKNSEDKPREGFTSTFGKYSSINDI-----DK 428
Db 472 RRIISSEGRAKIEMESSLRHKVQELLTVSK-FNDLKONDPTLAALCSTNDVLOQTDI 530
QY 429 IKKKK-----KGLIN-YKSTLYNNTINKKNNNNNNNNND 463
Db 531 VLQNTRAQLREEMRLCAHERTEHQLQDVGGLISTLGQVEDINSLSK-----LD 582

```


Db 221 NCEGSKVSHNLNVLADGSEBAQTAAGVRLK--EGCINRSLFLLQGVIKKLSDGQVGG 279
 Qy 283 HIFRDESLTVADLDFVGGKSKIMINISPTISCEBTANTLKYSSVVRPKRSTCIN 342
 Db 280 FINRDSLTNRLNLSLGNSEKTRILITTPV--SFDETLPALOPASTAKTKW--TPVNV 336
 Qy 343 E--DDPTNTERI--SILDSK-----GSENNASSIEN-----VIKSNHLNNNNNNKINR 388
 Db 337 EVSTDEALIKRYKEINHLKQLEVSLETPAQAMEKQALQLEBKLLQKQNEKEN 396
 Qy 389 -----GKINDKIERN-----N-399
 Db 397 LTRMLVTSSSLTLOOELFAKRRRVTCIGKINKKNSNVADQENIPNITTKTKHLKSLN 456
 Qy 400 ILK--NNSFDPKREGFTSPGKYSLNDIKIKKKKKGLINYSLTLYNDITINKKHN-- 456
 Db 457 LLREIDSEVCSSESVFNT-----LDTLSEIEMNPATKLNQENISELSLRADYDNL 510
 Qy 457 -----NN 490
 Db 511 VLDTYEQLTEKEEMELKKEKQDDDEPALKRKTQKQDEMQLIHISLTKLVGRVYVN 570
 Qy 491 --INNNTNN 537
 Db 571 ODLENELESSKVELLREKEDQIKLQEYIDSQLENIKMDLSYLSIESIEDPKQMKQTLFDA 630
 Qy 538 MNCMLNNNDKSIPLHKKRLRDNIKLNKSSCDNIMNKKNNLHLARHSVSGKLTMFSDP 597
 Db 631 ETVALDARESAFRLSENLELEKEKKEKELAT--TYKQEMNDIQLYQOLEKAKKK--QVLD 686
 Qy 598 QKNKDNTPFKSNINIKM--EDNTPKDIYE-----SRNVSMMNGNV 635
 Db 687 EKELOASAF--NEITLTJLIDGKVPKDLCLNLEBEGKTIDQKEINKEVEENALRESVI 744
 Qy 636 LLGLNKNTHDI-----STKOE-----NENDKKNNGVINIINNSNV 672
 Db 745 LLSEKSLPSSEVERLKEIODEKSELAHITTEKDLFSEVVKESRQGLLEIKGTKXD 804
 Qy 673 NSINNSNNNSINNSNNNSIYKSYNNNSISDVQIYV--VNMDSISNKNNDIIFPDA 729
 Db 805 LATTQSNKYSTDOEFONFKTLH-----MDERQRYKAVLEENEMODEIVALSKEA 854
 Qy 730 ISCDNNM-----YPNITNN 768
 Db 855 QKFPSSSLGALKTELSTYKTOLOEKREVOERLNEHQLEQLENDPSLQTYVEREKTILIT 914
 Qy 769 -----TNSMKLYAVNSHL-----FOPDNKNKTSNIQ--NINTKNKNDGNNVYSAN 814
 Db 915 EKLQOTLEEVKTLTQEKDQLKQLOESLQIERDQKSIDHTVNMNIDTQOQLRNALLESILK 974
 Qy 815 FCHYNLNDKNTVLIDLANKEQDKNIHGDDNIIQNRNDPEKKKTNFNANNVIYVNMN 874
 Db 975 QHOETINTLKSKI-----SEVSRLNH--MEENTEGETKDEQOKM----- 1012
 Qy 875 GNNNSPRMKGGLCGSHSIDMKNEMKNEMKDNEMKDNHKSNNNNSSSSSSNNNIY 934
 Db 1013 -----VGIDKKQDLEANTQTLTADVADNEIIEQOKRIFSLIQKELQ 1056
 Qy 935 NNI-----NDDTFQN-----DYCHNDTFTIRKNTNTINSNIYQ 970
 Db 1057 QMLESVIAEKEQLTKDLKENIEMTLENQBELRLLDDELKQOELIAQOENHA-----TKK 1111
 Qy 971 NDDIITYTINSLNDY-----MSNTLHKEKETYTYTSTNE-----DIYKE 1011
 Db 1112 EGEISRTCDRLAEVEKKEKSQOLOEQOQLLVNOEEMEQOKINIEINLKNLKEKE 1171
 Qy 1012 MEKGIH---RLDDQOKYDNDNNNDNNNNKNNVNDNNVNDNNNDNNNDKNNVND--- 1064
 Db 1172 LTLHEMERLELAKL-----NENVEVKSITERKVLKELQKQSFETERHLAGY 1222
 Qy 1065 -NNVD-----NDDVDVFNINIKNNNENYLSYFQKAVDTIINNCLNSLDISMY----- 1112
 Db 1223 IREIBATGLQTEKEELKIAHILHKEHOETIDELRBSVSEKTAQIINTQDEKSHTLQOEI 1282

Qy 1113 -----DOTEKILNNI--LTSKYAEKDNVIKYYI-----NE 1141
 Db 1283 PVLHEBEBLPPNKVSVSETQETMNELELLLEQSTTOSTTLARIEMERLNEKFOESQ 1342
 Qy 1142 DIKMS-----LEEDIKTAQSIYEKRVKVLTYLTLFFKKNVDTOI 1181
 Db 1343 EIKSLTERDNLTKEALEVHQDLQKEHIRETLAKQOESQKQOESLNNKEXNETTKI 1402
 Qy 1182 NNETS-----LRDVMCHICNNPPDQHFYAVSLBKDIINLIMROIWCESEN 1213
 Db 1403 VSEMEQFKPSQALLRIETIMLGLSKPLQSHDBMSVAEKDQLQLEVLO--SSDQ 1460
 Qy 1234 LRLTQFLAVEY-----ONKSANSVLINYSNNNDIILLNKL--VQDN 1275
 Db 1461 LKENIKETIVAKHLETEBELKVAHCCLEKQOEBETINELRVNLSKETELSTIQOLEALNDK 1520
 Qy 1276 IKNSMDHNNIHKK 1288
 Db 1521 LQNKI--QETIYEK 1531
 RESULT 21
 ID K168 DROME STANDARD; PRT; 784 AA.
 AC P46867; O9VTN8;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Kinesin-like protein KLP68D.
 GN KLP68D OR KLP5 OR CG7293.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachyceta; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI TaxID=7227;
 RN NCBI
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95050960; PubMed=7525600;
 RA Peravento P.A., Stewart R.J., Goldstein L.S.B.;
 RT "Characterization of the KLP68D kinesin-like protein in Drosophila;
 RT possible roles in axonal transport.";
 RL J. Cell Biol. 127:1041-1048 (1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer G., Champe M., Pfeiffer B.D.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Baau R.A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Boltsakov S.,
 RA Botkova D., Botchan M.R., Bouck J., Brocksen P., Brottier P.,
 RA Butcher K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davis P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.R., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dublin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fowler C., Gabriellian A.B., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harrie M.,
 RA Harris N.L., Harvey D., Helman J.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jajalji M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulj D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Nodarity C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusekern D.R., Paclob J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Rainer K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slater E., Stradling A.C., Stapleton M., Strong R., Sun E.,
 RA Stykasz R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasmann D.A., Weinstock G.M., Weissendach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "the genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP SEQUENCE OF 220-342 FROM N.A.
 RX MEDLINE=92020874; PubMed=1924306;
 RA Stewart R.J., Pesavento P.A., Woerpel D.N., Goldstein L.S.B.;
 RT "Identification and partial characterization of six members of the
 RT kinesin superfamily in *Drosophila*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8470-8474(1991).
 CC -1- FUNCTION: PLUS-END DIRECTED MICROTUBULE MOTOR THAT MAY BE USED FOR
 CC ANTROGRADE AXONAL TRANSPORT AND COULD CONCEIVABLY MOVE CARGOES IN
 CC FLY NEURONS DIFFERENT THAN THOSE MOVED BY KINESIN HEAVY CHAIN OR
 CC OTHER PLUS-END DIRECTED MOTORS.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PRIMARILY IN THE CENTRAL NERVOUS
 CC SYSTEM AND IN A SUBSET OF THE PERIPHERAL NERVOUS SYSTEM DURING
 CC EMBRYOGENESIS.
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
 CC IS SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 DR EMBL; U15974; AAA69929.1; -;
 DR EMBL; M74431; AAA28658.1; -;
 DR EMBL; AE003543; AAF50008.1; -;
 DR PIR; A55236; A55236.
 DR HSSP; P17119; 3KAR.
 DR FLYBase; FBgn0004381; Klp68D.
 DR GO; GO:0003774; F:Motor activity; IDA.
 DR GO; GO:0008089; P:anterograde axon cargo transport; IEP.
 DR InterPro; IPRO01752; kinesin_motor.
 DR Pfam; PF00225; kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
 KM Motor protein; Microtubules; ATP-binding; Coiled coil.
 FT DOMAIN 16 275 KINESIN-MOTOR.
 FT DOMAIN 351 385 COILED COIL (POTENTIAL).
 FT DOMAIN 426 582 COILED COIL (POTENTIAL).
 FT NP BIND 106 113 ATP (POTENTIAL).
 FT CONFLICT 220 221 SS -> TC (IN REF. 3).
 FT CONFLICT 338 342 GSRK -> VRGV (IN REF. 3).
 FT CONFLICT 338 338 G -> A (IN REF. 2).
 SQ SEQUENCE 784 AA; 88193 MW; 94BB9BAF072DFC0 CRC64;
 Query Match 7.0%; Score 477; DB 1; Length 784;
 Best Local Similarity 22.4%; Pred. No. 3,4e-14;
 Matches 185; Conservative 130; Mismatches 291; Indels 220; Gaps 27;
 QY 2 NSKIVVVRKRLPELEKKKSDITVKNKCTLYIDEPYKVDMTKYI-----ERHEPT 56
 DB 17 NECVVVVVRCRPMNSERSERSPEVVVYVN-----RGVELQNVVGNKEQKRVFT 68
 QY 57 VDCVFPDVTDFNVYVYENTIKELIIDLYE--NGCVSCFAFGCGSGKTYTMVLSGQPYGGS 114
 DB 69 YDAAYDASATQTTTLHEVVFPLVSVLEGNGLT---FAYGQGTGKTFTMBGV--GMD 123

QY 115 DPEGIFQYAGDIFTELNIYDKNTKG----IFISFYEYICGLYDLQKRKVAALENG 170
 DB 124 ELMGILPRFEQWILHIN-----RTENPQFLVDVSYLIEYMEELRDLKPNKSKHLEVRER 178
 QY 171 KKEVVKDKILKRLVLTKEELIKMIDGVLKRIQVNSQNDSSRSNAI--LNDLKDKNK 228
 DB 179 GSCVVPVNIHAINCKSEVEMIKMQGNKRRVGVGFNNHSHSRSHAIMIKILEMCDTET 238
 QY 229 NT-SLCKIAFIPLAGEGADTVSQNKQOTQDQANIRSLALKECIRAMDSKNIIPR 287
 DB 239 NTKVQKMLIDLAGESROSKTGASAEKAK-EASKINLALSLGNTVLSALBESSPHVPR 297
 QY 288 DSELTQVLDIFVQSKSIMANISPTISCCEQTNTLTYSRVKPKKSTCINEDDT 347
 DB 298 DSLTRTLQDSLGNSKTIIMINIGPSVNNVETLTLLAYGRASIQNP--IKNEDEQ 355
 QY 348 NT-----ERI-----SLDSKSEMASSIEN 369
 DB 356 DAKLKEYOERIRLRLIGPQOOQSEKQVAKKQVKKPKKRYTKEMSDSLQVSTIRQ 415
 QY 370 VTKSNH-----LLSNNNKINRGKINDKIERNNILKNSPKPRE-----GFTSTG 418
 DB 416 PVEDSDPEGASESDKENEAEVAKSNELEKERYNSKLAALAELEGOLYFGGNULD 475
 QY 419 KYSSLN-DIDK---IKNKKKGL-INYKSTLYNDNTINKGNNNNNNNNND----- 464
 DB 476 TYSERQIELEKTLVLAERKKREIEIQQLLEQETTLERENNVLSEQVELKRRKLG 535
 QY 465 -----NDNNNNNNNNND--SSMYNN-----MINMINNNNNNNNN 503
 DB 536 CVAKYLALQOEINDCKSDHNDLRELEMAQNELVYELKRLIIDFVIEVQRLYTOA 595
 QY 504 NNNNNNNNNNNHNNHPQNVYAFPTDSDPSLDDNNCHLNNNDKSLFLKQKULRDKIK 563
 DB 596 KIDEEDEKWKFSMSLPT----- 614
 QY 564 NRSSCDINKKKNNMLARHSVSKLTWFSYDPQKNDKTEPKSNINKEDNTPRDILY 623
 DB 615 --PGDGKRSKRPVSH-----PQRRRTSYA--LQAKNSPESSLAF 654
 QY 624 ESRNVSNNN-----GNVLLGNKTHHIDISTGDNHNDKINGV 663
 DB 655 KSENIYVLELPCPTTQERTPKVASLQAVLAQAMQAGDIDIVDS--HTNSLRSL 712
 QY 664 INIINN-----NVNSINNNSNNNSNNNSN 691
 DB 713 ENIINNANGAGPGAGVAGVSSIPVNRKIKSRGLPSAASWDSN 758
 RESULT 22
 KFP3B_HUMAN
 ID KFP3B_HUMAN STANDARD; PRT; 747 AA.
 AC 015066;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE kinesin-like protein KIF3B (Microtubule plus end-directed kinesin
 DE motor 3B) (HH0048).
 GN KIF3B OR KIAA0359.
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97349984; PubMed=9205841;
 RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. VII.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 4:141-150(1997).

[2]
 RN SEQUENCE FROM N.A.
 RA MEDLINE=2638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavriles G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clague S., Coley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Collison A., Coville G.J., Deacon R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Griffith D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lehar V.L., Martin S.L., Leverisa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McComachie L.J., McElay K., Murray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,
 RA Rice C.M., Ross M.T., Scott C.E., Senra H.K., Showkhen R., Sims S.,
 RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.,
 RT "The DNA sequence and comparative analysis of human chromosome 20.",
 RL Nature 414:865-871(2001).
 [3]
 RN IDENTIFICATION IN A COMPLEX WITH SMC3 AND KIFAP3B.
 RP MEDLINE=98175913; PubMed=9506951;
 RA Shimizu K., Shirataki H., Honda T., Minami S., Takai Y.,
 RT "Complex formation of SWAP/KAP3, a KIF3A/B ATPase motor-associated
 RL protein, with a human chromosome-associated polypeptide.",
 RL J. Biol. Chem. 273:6591-6594(1998).
 CC -I- FUNCTION: Involved in tethering the chromosomes to the spindle
 pole and in chromosome movement. Microtubule-based anterograde
 translocator for membranous organelles. Plus end-directed
 microtubule sliding activity in vitro (By similarity).
 CC -I- SUBUNIT: Heterodimer of KIF3A and KIF3B (By similarity). Interacts
 CC with the SMC3 subunit of the cohesin complex.
 CC -I- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
 CC II SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; AB02357; BAA20815.1; -
 DR EMBL; AL121897; CAC16425.1; -
 DR HSSP; P17119; 3KAR.
 DR GENE; HGNC:6320; KIF3B.
 DR MIM; 603754; -
 DR GO; GO:0005873; C:plus-end kinesin complex; TAS.
 DR GO; GO:0003777; F:microtubule motor activity; TAS.
 DR GO; GO:0008574; F:plus-end-directed kinesin ATPase activity; TAS.
 DR GO; GO:0008089; F:anterograde axon cargo transport; TAS.
 DR GO; GO:0007368; P:determination of left/right asymmetry; TAS.
 DR InterPro; IPR001752; kinesin_motor.
 DR Pfam; PF00225; kinesin; 1.
 DR PRINTS; SM00129; KINESINHEAVY.
 DR SMART; SM00129; KITS; 1.
 DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 DR PROSITE; PS0067; KINESIN_MOTOR_DOMAIN2; 1.
 KW Motor protein; Microtubules; ATP-binding; Coiled coil; Neuron.
 FT DOMAIN 1 345 KINESIN-MOTOR (BY SIMILARITY).
 FT DOMAIN 346 579 COILED COIL (BY SIMILARITY).

FT DOMAIN 580 747 GLOBULAR.
 FT NP_BIND 96 103 ATP (POTENTIAL).
 FT DOMAIN 386 393 POLY-GLY.
 FT DOMAIN 394 406 POLY-GLU.
 FT DOMAIN 723 730 POLY-SER.
 SQ SEQUENCE 747 AA; 85125 MW; 97FA4573AFA87023 CRC64;
 Query Match 6.9%; Score 472.5; DB 1; Length 747;
 Best Local Similarity 32.0%; Pred. No. 5.1e-14;
 Matches 136; Conservative 74; Mismatches 148; Indels 71; Gaps 14;
 QY 2 NSKIKVVRKRPSELEK-----KKQSDI-----ITYKNNCTLYIDPRKVDWTKIER 52
 DB 7 SESVAVVRRCRPMNGKEKASVDKVDVAVLGVSVKNRKGTAHEMPK----- 55
 QY 53 HEPIVDVFPDPTDNTFVVENTIKPLIIDLYE--NGCVGSCFPAVGGNGSKTYTMLGSOP 110
 DB 56 -TFEFDVAVPMNAKQFELVDETRPLVDSVLQNGTIT--FVGGTGKTYTMESGIR- 110
 QY 111 YGSDTGTGIFQYAAAGDIFTEFLNIYDKNTGIFISFEYICGKLYDLQK--RKMVALE 168
 DB 111 -GDEKRGVLPNSGDHIFTHIS-RSQNQVLVASVLEIYQEIIRDLSDQTKRLTKE 168
 QY 169 NGKKEVVVKDLKILRVLTKEELIKMDGVLRKIGVNSQNDSSSHALNTDK----- 224
 DB 169 RPDGVVVKDLSFVTSKVEIEHVMNVGNQNSVGTANNHSSRSRAIFVITIECEV 228
 QY 225 --DINKRTSGKIAFIDLAGSEGDADVSONKQOTQDGMNINSLAKECIRAM-DSDK 281
 DB 229 GLDENHIRVGKLNLDVLASGERAKTGAQGERLK-BATKINISLSLGNVISLVDGKS 287
 QY 282 NHIFRDESEKTVLRDIFVGSKSIMIANISPTISCCQTLNLRVSSRYKNFKSTCI 341
 DB 288 THIVRSKTLRLQDLSLGNNAKTVVAVNGPASYNEETLTLRVANRAKNIKNGR-V 346
 QY 342 NEEDDTERISLIDSGSENNASSINNVVIXNHLISNNNNKINRGKINDKIERNITL 401
 DB 347 NE-----DPDALRFEQEIARLKA-----QLEKRSIG 375
 QY 402 KNKSPDKREG 412
 DB 376 RKRKRERRREG 386
 RESULT 23
 KIF3B_MOUSE
 ID KIF3B_MOUSE STANDARD; PRT; 747 AA.
 AC Q61771;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Kinesin-like protein KIF3B (Microtubule plus end-directed kinesin
 DE motor 3B).
 DE KIF3B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBT_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR; TISSUE=Brain;
 RX MEDLINE=96032268; PubMed=7559760;
 RA Yamazaki H., Nakata T., Okada Y., Hirokawa N.,
 RT "KIF3A/B: a heterodimeric kinesin superfamily protein that works as a
 RT microtubule plus end-directed motor for membrane organelle
 RT transport.",
 RL J. Cell Biol. 130:1387-1399(1995).
 CC -I- FUNCTION: Involved in tethering the chromosomes to the spindle
 CC pole and in chromosome movement. Microtubule-based anterograde
 CC translocator for membranous organelles. Plus end-directed
 CC microtubule sliding activity in vitro (By similarity).
 CC -I- SUBUNIT: Interacts with the SMC3 subunit of the cohesin
 CC complex (By similarity). Heterodimer of KIF3A and KIF3B.

CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
 CC IF SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D26077; EMBL: A57107.1; --
 CC PIR: A57107; A57107.
 CC DR HSSP: P1119; 3KAR.
 CC DR MGDI: 107688; KIF3D.
 CC DR InterPro: IPR001752; kinesin_motor.
 CC DR Pfam: PF00225; kinesin_1.
 CC DR PRINTS: PR00380; KINESINHEAVY.
 CC DR SMART: SM00129; KISC; 1.
 CC DR PROSITE: PS00411; KINESIN MOTOR DOMAIN1; 1.
 CC DR PROSITE: PS00657; KINESIN MOTOR DOMAIN2; 1.
 CC KM Motor protein; Microtubules; ATP-binding; Coiled coil; Neutone.
 FT DOMAIN 1 345 KINESIN-MOTOR (BY SIMILARITY).
 FT DOMAIN 346 579 COILED COIL (BY SIMILARITY).
 FT DOMAIN 580 747 GLOBULAR.
 FT NP BIND 96 103 ATP (POTENTIAL).
 FT DOMAIN 386 393 POLY-GLY.
 FT DOMAIN 394 405 POLY-GLU.
 FT DOMAIN 723 730 POLY-SER.
 SQ SEQUENCE 747 AA; 85288 MW; FA369A4190EC8A47 CRC64;
 Query Match 6.9%; Score 472.5; DB 1; Length 747;
 Best Local Similarity 32.0%; Pred. No. 5, 1e-14;
 Matches 138; Conservative 74; Mismatches 148; Indels 71; Gaps 14;
 QY 2 NSKIKVVRKRPPLSELEK-----KKKDSI-----ITVKNCTIYDEPRYKVDTKYIER 52
 DB 7 SESEVRVVRRCRPMNGEKAKASYDKVDVAVLQGVSVKNGKTSHEMPK----- 55
 QY 53 HEPIVDKVPDDYDNTFVYENTIKPLIDLYE--NGVCSCFAYGQSGSKTYTMLGSP 110
 DB 56 -TTFPAVVYDWMNAKQFELVDYETFRPLVDVLAQGFNGTI---FAYGQGTGKTYTMEGRV- 110
 QY 111 YGQSDTPGLFQVYAGDIFTEFLINLYDKNTGKIFISFEIYCGKLYDLQK--RKMVALE 168
 DB 111 -GDEPKRGVTPNSFDHIFTHIS-RSQNQOVLVRSAYIEIYQERIDLSQDQTKRLLEKE 168
 QY 169 NGKEVVVVDKILRLVTKKELLKMDGVLLKRGIVNSQNDSSSRSHALINIDLK--- 224
 DB 169 RPDGVVYVDLSSFVTKSVKEIEHVNVVGNQNSVGTATNNHSSRSHALFVITIECEV 228
 QY 225 --DINKNTSLGKIAFLDLAGSERGADTVSONKOTQTDGAININSLAKKCIAM-DSDK 281
 DB 225 GLDGENHIRVGKILNVDLAGSEKQAGAEGLK-EATKINISLSMGVIALVDGKS 287
 QY 282 NHI:PFDSSELTXYLRIDFVSKSISIMANISPTISCEBOTLNTLRSSRYKNKSTCI 341
 DB 282 THIPYDSKTRILQDSLSGNATKTVANVAGPSVYVEELITLIRANRKAIKKMR-V 346
 QY 342 NEEDDTNTERISILDSKSEMANASSIENVVVKSNHLLSNNNKINKNGKINDKIERNMIL 401
 DB 347 NE-----DPKDLALREFQEIARLKA-----OLEKRSIG 375
 QY 402 KNSKPKPKREG 412
 DB 376 RRRRRRKRREG 386

RESULT 24
 KLP3_SCHPO
 ID KLP3_SCHPO STANDARD; PRT; 554 AA.
 AC Q9US60; Q9US61;
 DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Kinesin-like protein 3 (Kinesin-related protein 1).
 GN KRP3 OR KRP1 OR SPAC1834.07
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetaceae;
 OC NCBI_TaxID=4896;
 RX SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.
 RC STRAIN=972;
 RC MEDLINE=22127603; PubMed=10641037;
 RT Brazer S.-C.W., Williams H.P., Chappell T.G., Cande W.Z.;
 RL Yeast 16:143-166(2000).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RC MEDLINE=22127603; PubMed=12132578;
 RT Jeong J.W., Rhee D.K., Cho S.Y., Hae K.L., Kim D.U., Won M., Kim H.B.;
 RL "Cloning and characterization of the kinesin-related protein, Krip3,
 in Schizosaccharomyces pombe.";
 RL Mol. Cells 13:389-398(2002).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RC MEDLINE=21848401; PubMed=11859360;
 RX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C.,
 RA Oliver K., O'Neill S., Pearson D., Quayl M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Wolcraets G., Aert R., Robben J., Grynopre B.,
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel K., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wambut R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revela J.L., Moreno S., Armstrong J., Forburg S.L.,
 RA Cernetti L., Lowe T., McCombie W.R., Paulsen O., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: CYTOPLASMIC MOTOR THAT COULD PLAY A ROLE IN GOLGI
 CC MEMBRANE RECYCLING.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF154055; AAF14525.1; --
 CC DR EMBL: AF156965; AAF22609.1; --
 CC DR EMBL: AF247188; AAF81205.1; --
 CC DR EMBL: AL157734; CAB75775.1; --
 CC DR PIR: T50118; T50118.
 CC HSSP: P33176; 1BG2.


```

Db 75 ITVAVRGRGRNREISKSSVAVVVPDITSGEISINTT-----GDTGTTAQMNKRY 127
Qy 56 IYDVFDYDNTFYENTTKPLIIDYENGCVCSFAYGOTSGKTYTMGOPY--GQ 113
Db 128 TVDKVFGPGASQDIDFVAVGPLEFQD--FIKGYNCTVAVYGMTSTGKTYMTGDEKLYNGE 186
Qy 114 -SDTPGIFQYAGGIFFTFLNYDKDNKGFISFYEYICGLVY----- 156
Db 187 LSDAAGIIPVLLKLPTELQNDYV--VKCSFTELNEBELKOLLDSNSGSGNTGPDG 244
Qy 157 -----LQKRKRVVALENGX 171
Db 245 QFMKKLIIFDSTANTNTSSASSRSRSRSPSLNDLTPKAAALRKRLRTSLPNTI 304
Qy 172 KE-----VVVKDLKLRV 184
Db 305 KOQYQOQOAVNSRNNSSSGSTNNASSTNTNGGSSWAPNDQNGIYOMQOEHI 364
Qy 185 LTKBELIKMIDGVLRLKIGVNSQNDSSRSHATLIDL--KDINKTSLGKIAPIDLAG 242
Db 365 TNAMGLNLQKGLKHROVASTKMNDFSRSHTFTTLTKKHQDELFRISKMLVLDLAG 424
Qy 243 SERGADTVSONKOTODGANINRSLALKECIRAMDSDKNHI PRDSELTFLVRLDIFVGX 302
Db 425 SENINRSGALNQRKAG--SINGSLTLGRVYNALVDKSGHLPRESKTLRLDLSLGN 483
Qy 303 SKSIMANISPTISCCQTLNTRYSSVFNKFKKSTCINEEDDTNTERISILDSKSEM 362
Db 484 TKTALITATISPAKVSEETCSLTVASAKNINKK-----POL 521
Qy 363 NASSIENVVKSNNHLSNNNNKINRGKINKIEBNNLKNKSPDKREGFTSGFKSS 422
Db 522 GSPIMKDLV-----NITMELAKI--KSDLSLTKS--KEGYMSODHYK 563
Qy 423 LN-DIDKX-----KNNKGLINYKSTLYNDNTINKKNNNNNNNNNNNNNNNNNN 475
Db 564 LNSDLESYKNEVOEKREIESLTSKNALVMDKLSKETISOQOIES----- 612
Qy 476 NNDSSVNNNNINEM--INNINNINIVNNNNNNNNNNNNNNNNNNNNNNNNNN 532
Db 613 -----LKTIDHILAOQDKHTEIISDFNNKLQK-----LFEVWQM 650
Qy 533 SSLDMNCHLNNNDK--SIFLHKKNLNDNLKLN--NRSSCONIMKK--KNNLHARSY 586
Db 651 ALHDKKRELDNQFEMWHITKELKSTFLQNLTMQOESILOETNIQPLDMTKEV 710
Qy 587 GSKLTMFSYDPOKKNKNTFFKSNINKMEDNTPKDIYESRVNSNNGVLLGLNKNTHD 646
Db 711 ---LTLNR--TMQEKALMYDCVAKILNESPK----- 738
Qy 647 ISTDENHNKINNGVNIINNSVNSINNSNNSINNSNNSIYKSNVNSQISD 706
Db 739 -----FVVV-----IEKIDIRVDFOKFYK--NIAMENISD 767
Qy 707 VOIRVNMEDTNSKND-----NIFPDAISCDNNMYRNTNNNNNNNNNNNNNNNDV 760
Db 768 I-----SENNNNKKOYLKXHP-----KNNHDELRLRHVDSTYENKKTNEFV 811
Qy 761 ENYNRRDGTNSMKLYAVYNSHLFOPDNKNTSNIQNTNKNQDGNVYEMNFCNYL 820
Db 812 ENFKK-----VANDH--LDEKIKLI--MQLTLTAS--AVIDQMDL----- 847
Qy 821 NDKNYLIDLNKEQDKN-----IHGDNNIIONRNDFEKKKKTFFYNNNNIYIVANNNGN 876
Db 848 -----FEPKRVKXWNSFDLINDCS--MNNIEFYNSMATLISQIKSTYDTSNSMN 895
Qy 877 NNSPRMKYGLGSHSISDMKNKMKNNKMKNNKMKNNKMKNNKMKNNKMKNNKMKNN 932
Db 896 BSISVWKGVSEENASILKNNKYFNDFEQLINKHNLKONIKSITSTSHSTHNTVD 955
Qy 933 IYNNINDDTFQNDYCHNDNTFTIRKNNNTINSNIYQDDDIYITNSINDYMSNTLLHF 992
Db 956 IYNTIE-----NIMKN--YGNKE-----NAKMDMIENIL-- 983

```

```

Qy 993 KEKYTPLTSTNEDIYKMEGKIRLDDQDKYDNDNNNNVNNNNKNNV-----DN 1043
Db 984 KE--IPNLS-----KKMP--LKL-----SNINSNVQSVISPKKHAIEDE 1019
Qy 1044 NVDNNNVND 1052
Db 1020 NKSENVND 1028

RESULT 27
KF3A HUMAN STANDARD; PRT; 702 AA.
ID KF3A HUMAN
AC 09496;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Kinesin-like protein KIF3A (Microtubule plus end-directed kinesin motor 3A).
DE motor 3A).
GN KIF3A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=20018159; PubMed=10548469;
RA Whitehead J.L., Wang S.Y., Best-Usinger L., Hoang E., Frazer K.A.,
RA Burnside B.;
RT "Photoreceptor localization of the KIF3A and KIF3B subunits of the
RT heterotrimeric microtubule motor kinesin II in vertebrate retina.";
RL Exp. Eye Res. 69:491-503(1999).
CC -!- FUNCTION: MICROTUBULE-BASED ANTEROGRADE TRANSLOCATOR FOR
CC MEMBRANOUS ORGANELLES. PLUS END-DIRECTED MICROTUBULE SLIDING
CC ACTIVITY IN VITRO.
CC -!- SUBUNIT: HETERODIMER OF KIF3A AND KIF3B.
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC II SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF041853; AAC72294.1; -.
DR HSSP; P17119; 3KAR.
DR Genew; HGNC:6319; KIF3A.
DR MIM; 604683; -.
DR GO; GO:0006996; P:organelle organization and biogenesis; TMS.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil; Neutone.
FT DOMAIN 1 350 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 351 587 COILED COIL (BY SIMILARITY).
FT NP_BIND 588 702 GLOBULAR
FT NP_BIND 100 107 ATP (BY SIMILARITY).
FT DOMAIN 443 446 POLY-GLU.
SQ SEQUENCE 702 AA; 80337 MM; 24507DIE6D540DB CRC64;

Query Match 6.7%; Score 460; DB 1; Length 702;
Best local Similarity 29.8%; Pred. No. 1.7e-11;
Matches 147; Conservative 84; Mismatches 192; Indels 70; Gaps 15;
5 IKVVRKRLSELEKKKKSDIITVKNKNTLYIDEPKYKDMTYIERHE---FYDVKV 60

```


Db 15 VKVVRRCRLNERKSKCYQAVS-----DEMGTTVHKHTDSSNEPKTFEDTV 66
 Qy 61 FDDTVDFVYENTIKPLILDYE--NGCVSCFAYGQSGSKTYTMGSPYQSDPTG 118
 Db 67 FGPSKQDLYNNLTARPIIDSVLENGNTI---FAYGQTGKGTFTMEGRAT--PELRG 121
 Qy 119 IFQYAADIFTEFLNIYKDNKTGIFISFYEIYCGKLYDLQK--RKNVAALENGKEVVV 176
 Db 122 IIPNSFAHIFGHIAKEGDTFLVAVSYLIEYNEEVLDLQDQTRLAVERPQGVYI 181
 Qy 177 KDLKILAVLTKKEELIKMIDGVLRKIGVNSQNDSSRSHILNIDK-----DINKT 230
 Db 182 KDLAAYVNNADDDRTMTLTGHKRRSGATNMNHSRSHIFITTECESEKIDGMHV 241
 Qy 231 SLGKIATIDLAGSRGADTVSONKOTODGANINRSLLATECIRAM--DSPKNIIPRDS 289
 Db 242 RMGLTHLVDLAGSRQAKTATGQRK--EARKILSLSTLGNVSLAVDGSTHVPFRNS 300
 Qy 290 ELTKVLRDIFVGSKSMIANISPTISCEQTLNLTARYSRKVNKPKSTCINEED--- 345
 Db 301 KLTFLRLDQSLGNSKTMWCANIGPADYNYDETISTLAYNARAKNIKKAR--INEDPKAL 359
 Qy 346 ----DITTERSLIDSQSEMMASSINENVYKSHLNNNNKINKIND----KIE 396
 Db 360 LRQFQKEIEELKKLEEGEISGSDI-----SGSEEDDEGEVEDEDEKRRK 408
 Qy 397 RNNILKKK-----SFDKPEGEFTSTFG-----KYSLNDIDIKIKKKKKGLNY 440
 Db 409 RIQGGKKVSPDKKIEOMAKIDERRKALETKLDMEBERKAKARALEKREKDLKAOEH 468
 Qy 441 KSTLYNDNTINKK 453
 Db 469 QSLLEKLSALEKK 481
 RESULT 28
 KL61_DROME STANDARD; PRT; 1066 AA.
 ID KL61_DROME
 AC P46863; OST0A6; OSW018;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Bipolar kinesin KRP-130 (Kinesin-like protein Klp61F).
 GN KLP61F OR CG9191.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=94043448; PubMed=8227131;
 RA Heck M.M.S., Pereira A., Pesavento P.A., Yannoni Y., Spradling A.C.,
 RA Goldstein L.S.B.;
 RT "The kinesin-like protein Klp61F is essential for mitosis in
 RT Drosophila";
 RL J. Cell Biol. 123:665-679(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Anagnostis P.G., Scherer S.E., Li P.W., Hopkins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintinas D.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Domes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.U., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foeller C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Goelzer A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Houston D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostreli A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Sylvestre R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP REVISIONS.
 RC STRAIN=Berkley;
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Beccencourt B.R., Celniker S.E., de Grey A.D.N.V., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley; TISSUE=Embryo;
 RX MEDLINE=22426066; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
 RA George R.A., Guarin H., Krommiller B., Paclob J.M., Park S., Wan K.H.,
 RA Rubin G.M., Celniker S.E.;
 RT "A Drosophila full-length cDNA resource";
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 RN [5]
 RP SEQUENCE OF 228-357 FROM N.A.
 RC STRAIN=DP CN BW;
 RX MEDLINE=92020874; PubMed=1924306;
 RA Stewart R.J., Pesavento P.A., Woerpel D.N., Goldstein L.S.B.;
 RT "Identification and partial characterization of six members of the
 RT kinesin superfamily in Drosophila";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8470-8474(1991).
 RN [6]
 RP PARTIAL SEQUENCE, AND IDENTIFICATION AS KRP-130.
 RX MEDLINE=97078747; PubMed=8918872;
 RA Keshna A.S., Scholey J.M., Leszyk J.D., Saxton W.M.;
 RT "An essential bipolar mitotic motor";
 RL Nature 384:225-225(1996).
 CC -1- FUNCTION: IMPORTANT ROLE IN MITOTIC DIVIDING CELLS. MICROTUBULE
 CC MOTOR REQUIRED FOR SPINDLE BODY SEPARATION. SLOW PLUS-END DIRECTED
 CC MICROTUBULE MOTOR CAPABLE OF CROSS-LINKING AND SLIDING APART
 CC ANTIPARALLEL MICROTUBULES. THEREBY PUSHING APART THE ASSOCIATED
 CC SPINDLE POLES DURING SPINDLE ASSEMBLY AND FUNCTION.
 CC -1- SUBUNIT: Homodimer.
 CC -1- DEVELOPMENTAL STAGE: SPECIFICALLY EXPRESSED IN PROLIFERATING
 CC TISSUES DURING EMBRYONIC AND LARVAL DEVELOPMENT.
 CC -1- PTM: PHOSPHORYLATION DURING MITOSIS AT THR-933 CONTROLS THE


```

CC      ASSOCIATION OF KL6P1F WITH THE SPINDLE APPARATUS (BY SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BINC
CC      SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      at the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by, and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U01842; AAA03718.1; -.
DR      EMBL; AE003471; AAF47458.2; -.
DR      EMBL; AY069442; AAL39587.1; -.
DR      EMBL; M74428; AAA28655.1; -.
DR      PIR; A48669; A48669.
DR      HSP; P17119; 3KAR.
DR      FLYBase; FBgn0004378; KLP61F.
DR      GO; GO:0005737; C:cytoplasm; IDA.
DR      GO; GO:0005871; C:kinesin complex; IDA.
DR      GO; GO:0003774; F:motor activity; IDA.
DR      GO; GO:0007100; P:centrosome separation; IGI.
DR      InterPro; IPR001752; kinesin_motor.
DR      Pfam; PF00225; kinesin_1.
DR      PRINTS; PR00380; KINESINHEAVY.
DR      SMART; SM00129; KISC; 1.
DR      PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR      PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
DR      Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
DR      Mitosis; Cell cycle; Phosphorylation.
FT      DOMAIN 17 361 KINESIN-MOTOR (BY SIMILARITY).
FT      DOMAIN 362 462 COILED COIL (POTENTIAL).
FT      DOMAIN 463 569 COILED COIL (POTENTIAL).
FT      DOMAIN 570 678 COILED COIL (POTENTIAL).
FT      DOMAIN 679 787 COILED COIL (POTENTIAL).
FT      DOMAIN 788 896 COILED COIL (POTENTIAL).
FT      NP_BIND 103 110 ATP (POTENTIAL).
FT      MOD_RES 933 933 PHOSPHORYLATION (BY CDC2) (BY
FT      SIMILARITY).
FT      VARIANT 595 595 M -> V.
FT      VARIANT 869 869 R -> K.
FT      VARIANT 904 904 L -> Q (IN REF. 1).
FT      CONFLICT 962 962 L -> Q (IN REF. 1).
FT      CONFLICT 983 983 V -> D (IN REF. 1).
SQ      SEQUENCE 1066 AA, 121163 MW, 363647366E0721F CRC64;

Query Match 6.7%; Score 459; DB 1; Length 1066;
Best Local Similarity 21.0%; Pred. No. 2.9e-13;
Matches 180; Conservative 175; Mismatches 327; Indels 176; Gaps 28;

QY      2 NSKIVVRKRPRLSELEKKKDSIIYKNNCTLYIDBRVAVMTKTIERH----- 53
DB      17 NONTQVVRVPLNRSRRCIRSAEAVDV-----VGPREVTRHTLDSKLTK 62
QY      54 EPIYDKVDDVNTVYENTIKPILIDYENGCVSCFAGQSGSKTYTMLGSG----- 109
DB      63 KFTFPRSGPESKQCDVSVVSPLEEVL-NGVNCYFAGQGTGKHTWVGEMLEL 121
QY      110 --PYQSDTPGIFQYAGDIFFTLNIYDKDNTKGFISFVEIYCKLYDL--OKRNV 164
DB      122 KSSWEDDDIGIIPRALSHLPDELMMVEYLT--KRISLELYNELCDLSTDDTTKIR 179
QY      165 AALENGKK-EVVVKDKILRVITKEELLIKMIDVGLRKIGVNSONDESSRSHALINTDL 223
DB      180 IFDDSTKGSVLIQGLEIPIVHAKSDVYVLLKLEKGERKRTATTLMAAGSSRSHVFSI-V 238
QY      224 KDINK-----TSLGKAFIDLAGSEKADTVSQNKTOTDGNINRSILALKECTRA 276
DB      239 VHIRENGJEGEDMLKIGKLNLDLAGSENVSKAGNEKIGRVREYVAINOSLLTLRLVITA 298
QY      277 MDSDKNHPFRDSELTQVLKRDIFVSKSKSIMIANISPTISCEQTLNLTFRSSRVNFKN 336

```

```

DB      299 LVDRAPVPRSEKTLRLLOESIGRTKTSIIATISPGHDIETLSTLEVAPRAKNION 358
QY      337 -----KSTCNER-----DDNTERISIIILSKSENMASSENIVIKSNLLSNNN- 382
DB      359 KPEVNOQLTKTKVLTKEETEEIDKLRDLMARDNGVLYLEETGEITLK-----LESGNRE 415
QY      383 -NNKI-----NRGINDKIERNNILKNKSPDKPREGFTSPFGKY-----SLINDI 426
DB      416 LNEKMLLLKALKEDELQKEKIFSEVSMVLKEQELKTEENLNTGTLTLTKVLTXT 475
QY      427 DKIKNNKKGLINYKST-----LYNDNT-----INKKA 454
DB      476 KRYEKEKELVASHMKTEQVLTQAOEILAAADLATDPTQLHGTTERREDEKIRSC 535
QY      455 NNNNNNNNDNNNDNNNNNNNNNNSSSMVNMIMNINN-INNNVNNNNN----- 506
DB      536 DQFKRDNDLEMIIGSLNIYQDQALKEQLSGEWMVSSVSGRLNLSKSIEMLEKEM 595
QY      507 ---NNNNNNNNHNNHLPQPVYAFTDTSFSLDDMCHLNNDKSIPLAKKNLRDNIKL- 562
DB      596 CAOSLODQTLNHLNKLGE-----VVKISDQHSQAFVAKLMQWQOQLMSKEIQTNLQVI 651
QY      563 ---KNRSSCDNIMNK-----KKNNLHLARHSVSKLTMTFSYDPQKXKN 603
DB      652 EENNQRHKAMLDSSQEFATIIISLSQSVBEHAKQMKLEQAGMSLPABEIQNLQEE 711
QY      604 TFEKSNINKMEDNTPKDIIVESRNVSMNG-----NVLTGLNKQTHHDISTDENHNDKI 659
DB      712 LANRALAQDQDALLBSMMQMEQIKVLRSKNSISMSVHLAK--MEESRLTRNHRIDDI 768
QY      660 NNGV-----INIINNSVNS-----INNSNNNSINNSNNNSIYSKNY 698
DB      769 KSGIQDYQKIGIEHSGAQAELISQMEAGMLCDQGVANQSMQVHKKNLNQ-----KYEX 824
QY      699 NSNGSIDVQIRYVNEKD 716
DB      825 ETNENVSQVVRVH-NQVE 841

RESULT 29
KIN_HUMAN STANDARD; PRT; 963 AA.
AC P33176;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE kinesin heavy chain (ubiquitous kinesin heavy chain) (UKHC).
GN KIF5B OR KNS1 OR KNS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=92299683; PubMed=1607388;
RA Navone F., Niclas J., Hom-Booher N., Sparks L., Bernstein H.D.,
RT "Cloning and expression of a human kinesin heavy chain gene:
RT interaction of the COOH-terminal domain with cytoplasmic microtubules
RT in transfected CV-1 cells."
RL J. Cell Biol. 117:1263-1275 (1992).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 1-349.
RX MEDLINE=96195066; PubMed=8606779;
RA Kull F.J., Sablin E.P., Lau R., Fletcher R.J., Vale R.D.;
RT "Crystal structure of the kinesin motor domain reveals a structural
RT similarity to myosin."
RL Nature 380:550-555 (1996).
RN [3]
RP TISSUE SPECIFICITY.
RX MEDLINE=94242426; PubMed=7514426;
RA Niclas J., Navone F., Hom-Booher N., Vale R.D.;

```


RT "Cloning and localization of a conventional kinesin motor expressed
 RT exclusively in neurons.";
 RL Neuron 12:1059-1072(1994).
 CC -1- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
 CC PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT.
 CC -1- SUBUNIT: OLIGOMER COMPOSED OF TWO HEAVY CHAINS AND TWO LIGHT
 CC CHAINS.
 CC -1- SUBCELLULAR LOCATION: UNIFORMLY DISTRIBUTED BETWEEN THE CELL BODY
 CC AND THE PROCESSES IN THE NEURONS.
 CC -1- TISSUE SPECIFICITY: FOUND IN NEURON AND ADULT BRAIN, LIVER,
 CC KIDNEY, SPLEEN, HEART, LUNG AND SCIATIC NERVE.
 CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS: A LARGE GLOBULAR N-
 CC TERMINAL DOMAIN WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY OF
 CC KINESIN (IT HYDROLYZES ATP AND BINDS MICROTUBULE), A CENTRAL
 CC ALPHA-HELICAL COILED COIL DOMAIN THAT MEDIATES THE HEAVY CHAIN
 CC DIMERIZATION; AND A SMALL GLOBULAR C-TERMINAL DOMAIN WHICH
 CC INTERACTS WITH OTHER PROTEINS (SUCH AS THE KINESIN LIGHT CHAINS),
 CC VESICLES AND MEMBRANOUS ORGANELLES.
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; X65873; CAA46703.1; -.
 DR PIR; A41919; A41919.
 DR PDB; 1BG2; 14-OCT-98.
 DR Genew; HGNC:6324; KIF5B.
 DR MIM; 602809; -.
 DR GO; GO:0005871; C:kinesin complex; TAS.
 DR GO; GO:0004002; F:adenosinetriphosphatase activity; TAS.
 DR GO; GO:0003777; F:microtubule motor activity; TAS.
 DR GO; GO:0007018; F:microtubule-based movement; TAS.
 DR InterPro; IPR001752; Kinesin_motor.
 DR Pfam; PF00225; kinesin_1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN; 1.
 DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
 KM Motor protein; Microtubules; ATP-binding; Coiled coil; 3D-structure.
 FT DOMAIN 1 256 KINESIN-MOTOR.
 FT NP_BIND 915 963 COILED COIL.
 FT STRAND 8 15 GLOBULAR.
 FT HELIX 20 25 ATP (BY SIMILARITY).
 FT STRAND 26 26
 FT STRAND 29 29
 FT STRAND 32 34
 FT STRAND 35 37
 FT STRAND 38 41
 FT TURN 42 43
 FT STRAND 44 47
 FT STRAND 50 52
 FT TURN 54 55
 FT HELIX 58 65
 FT HELIX 66 66
 FT TURN 67 74
 FT TURN 75 76
 FT STRAND 77 84
 FT TURN 87 88
 FT HELIX 91 95
 FT TURN 96 96
 FT STRAND 97 97
 FT TURN 99 100
 FT TURN 102 104
 FT STRAND 105 105
 FT HELIX 107 122

FT STRAND 126 138
 FT TURN 139 140
 FT STRAND 141 144
 FT TURN 148 149
 FT STRAND 153 153
 FT STRAND 155 157
 FT TURN 159 160
 FT STRAND 163 165
 FT TURN 166 167
 FT STRAND 171 173
 FT HELIX 176 189
 FT TURN 190 193
 FT HELIX 197 203
 FT STRAND 205 216
 FT TURN 217 219
 FT STRAND 222 231
 FT TURN 246 247
 FT HELIX 256 269
 FT TURN 270 271
 FT HELIX 277 279
 FT HELIX 281 285
 FT STRAND 286 288
 FT TURN 289 289
 FT STRAND 295 302
 FT STRAND 305 305
 FT HELIX 306 308
 FT HELIX 309 320
 FT TURN 321 321
 SQ SEQUENCE 963 AA; 109685 MW; A1PE5760C3250C8B CRC64;

Query Match 6.6%; Score 453; DB 1; Length 963;
 Best Local Similarity 21.8%; Pred. No. 4,7e-13;
 Matches 215; Conservative 171; Mismatches 421; Indels 178; Gaps 34;

QY 5 IKVVRKRPISLEKKKKKSDIITVKNKCTLYIDEPKRYKMTKTYIERHEFIDKVPDDT 64
 9 IKMCHFRPLINESEVNRGDK-----YI-AKQGEDTVIASKPYAFDRVQSS 55
 QY 65 VDNFTYENTIKLIIDLYE--NGCVSCFAYGOTSGKTYMTLGS--QPYGSDPDPGF 120
 56 TSOEYVNDCAKIVDQVLEGVNGTI--FAYQTSKGKHTMEGLHDEGNGIIPRIV 112
 QY 121 QYAAQIFPEFLNYDKDKNTGIFISFEYICGLYDLQKRKVAALENGKKEV-VVKDL 179
 113 Q-----DIFNYIYMDENLEPHIKVSYFEIYLDKIRDLDSKTNLSVHEDKRVPYVKG 168
 QY 180 KILRVLTKEELIKMIDGVLARKIGVNSQNDSSSRSHAILNIDKIDINKTS---LGKIA 236
 169 TERFVCSPEVMDTIDEGKSNRHVAVTNNNEHSSRSHIFLINVKQENTQTEQKSGKLY 228
 QY 237 FIDLASSEKADPVSONKOTQTDGANINRSLALKECTIRAMDSKNIHPPRDELTQVLR 296
 229 LVDLASSEKSVSKTGAGG-AVLDEAKNINSLALGVNIALAGSTYVYPRSKMRILQ 287
 QY 297 DIFVGSKSIIMIANIPTISCCBOTLTLRYSRRVNFKNKSTCIINEEDPTNTERISILD 356
 288 DSLGNCRTTIIVCCPSSSYNESETSTILFGRATIKK-TYCAVVE----- 334
 QY 357 SKGSENNASSIENVVIKSHLLSNNNNNKINRGKINDKIERNNI--LKNGKSPDKPRGFT 414
 335 -----LTAEQWKKKYKEKEKNIKILRNITQMLENE-LNRMRNGET 373
 QY 415 STFGKYSLSNDIDKIKKNGKGLINYSKTLVNDNTINKGHNNNNNNNNNDND----- 466
 374 VPIDE-----QDFKEVLANLEAFVTDKITLTDKPATYAGVIGNFTDARRKCEEBIATKL 428
 QY 467 -NNNDNNNNNNNDSSGVANNIMINHMNNNNNNNNNNNNNNNNNNNNNNNNHNLPOPNYA 525
 429 YKQLDQKDEINQSGVLEKTKQMLDQF-----ELLASTRDQDMQALNRLQLENDA 483
 QY 526 PTD--TSDFSSLDNNKCHLNNDKSTFLHKKNLRDNIKLNKSSCDNINAKKKKNINILHAR 583
 484 SKEEVEKVALLEEL--AVNYDQKS-----QEVED--KTKEYELASDELINQKSATL---- 530
 DB

DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Kinesin-like protein 5.
 GN KLP5 OR SPBC2F12.13.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetes; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 CX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
 RX MEDLINE=21963728; PubMed=11967147;
 RA Garcia M.A., Koonruga N., Toda T.;
 RT "Two kinesin-like kin I family proteins in fission yeast regulate the
 establishment of metaphase and the onset of anaphase A.";
 RL Curr. Biol. 12:610-621(2002).
 RL [2]
 RN SEQUENCE FROM N.A.
 RX STRAIN=972;
 RA MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooke K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odel C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovich E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Seaton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voicakeert G., Aert R., Robben J., Grymoprez B.,
 RA Weltjens I., Vansteede E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Janger I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rocher M., Gallard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito G.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Sipkavski G.V., Ussery D., Barrett B.G., Nuree P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -!- FUNCTION: Has a role in establishing metaphase during mitosis.
 CC Required for chromosome segregation where it generates tension
 CC during kinetochore capturing.
 CC -!- SUBUNIT: Heterodimer with klp6.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic microtubules in interphase,
 CC mitotic kinetochores in metaphase and spindle midzone in anaphase
 CC and telophase.
 CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
 CC II SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1ab-sib.ch/announce/>
 CC or send an email to license@1ab-sib.ch).
 CC -----
 DR EMBL, AB072924, BAB69885.1; -.
 DR EMBL, Z97211, CAB10160.1; -.
 DR PIR, T40128; T40128.
 DR HSP, P17119; 3KAR.
 DR GeneDB Spombe: SPBC2F12.13; -.
 DR InterPro: IPR001752; kinesin_motor.
 DR Pfam, PF00225; kinesin_1.
 DR PRINTS, PR00380; KINESINHEAVY.
 DR SMART, SM00129; Kisc; 1.

DR PROSITE, PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE, PS0067; KINESIN MOTOR DOMAIN2; 1.
 KM Mitosis; Chromosome partition; Motor protein; Microtubules;
 KW ATP-binding; Coiled coil.
 FT DOMAIN 1 388 KINESIN-MOTOR (BY SIMILARITY).
 FT DOMAIN 396 435 COILED COIL (POTENTIAL).
 FT DOMAIN 563 588 COILED COIL (POTENTIAL).
 FT NP BIND 144 151 ATP (POTENTIAL).
 SQ SEQUENCE 883 AA; 99058 MW; BB8A60E22DA397E8 CRC64;
 Query Match 6.5%; Score 449; DB 1; Length 883;
 Best Local Similarity 21.3%; Pred. No. 6,4e-13;
 Matches . 218; Conservative 159; Mismatches 354; Indels 294; Gaps 37;
 3 SKIKVVRKRPISLEKKK-KDSD-----IITVKNCTYIDE 39
 5 SSTIVTVRRPFTASANLIASDLSIGTSSSLNPGSGRIKRVVYLDKRVLVDP 64
 40 PRYKVMTKYIEH-----EFIVKVPDDTVNFTYEYNT 74
 65 PDETATVLSATNRLSTSQOGLRLSRKSNNSAGFGRDLRYADRVDEFATQOQYERT 124
 75 IKPLIIDLYENGVCSCFAYGQSGKTYTMLASQYGGSDTTCIROYAAGDIFTELVNT 134
 125 ARPLDNLID-GFNATIFAYGATGCGKTHISGTM-----ODPGLYLTLEKLFERMDHL 178
 135 DKONTGIFISFEIYCGKLYDLDL-----QKRMVALENGKKEVVVKDLIRVLTKEE 189
 179 RDKRTIDRLSYELINETIRDLVSPTRQAKPLVREDADRRITVPGLTISPSLEE 238
 190 LILKMDIGVLRKIGVSONDESSRSHAILNIDL-----KINKNTSGKIAFIDLAGS 243
 239 IIDIIKGNANRMTSPTEANASRSHAVLQVTLQKPRTAGINEDHTLATISIDLAGS 298
 244 ERADTVSQKQQTGTGAININSLAKECIRAM--DSQKNIPFDSLETUKVLPFG 301
 299 ERATATKLGSRF-FEGAININSLALGNCINALCPHRAHVPYDSKTRLLKESLGG 357
 302 KKSINMIANISPTSCCEOTLNTLRSSRVKPKNKGSTINEEDDINTERISIDSGSE 361
 358 NCRIVAVCVSPSVYEEHTNLKYANPAKNIKT-----EYLRNNTISV-----D 402
 362 MNASILENVVIVKSNHLLSNNNNNKIRKINDKIEBNILKNKSPDREGFTSTFGKYS 421
 403 RHVSQVKAIVELREGISELRILA-----QID-----LSQSGNSDDQAVTQSAHRS 451
 422 SLNDIDIKIKKKKKGLINTKSTL-YNDNTINK--KANNNNNNNDNDNNNNNNNN 476
 452 KLAENRNLRL-----MFEETLPLQNDTINKVEKYKH-----F 484
 477 NQSSSVNMMINMINNN 536
 485 DDSIRVAKTQVL-----SCYERILP----- 503
 537 DMNCHLNNDKSIFFLKKKYL-----DNIKLNKSSCDNIMNKKNNLHL 581
 504 -----NSADERVFVLRSGLESLTTRARILADIDELVYQKFRQSVSHIINTYKQ----- 553
 582 ARHVSQKLTWPSYDPQKNKNTFFKSNL-----NKQEDNTP--KOILYESRVNSNM 631
 554 -----EGATVYADVLDQDQVD--LKSIITENQVLDQKNYVDEPTPLVESILSSPFASSIL 605
 632 -NGNV-----LGLNKNTHDITQDENHNDKINGVINIINNSVNSISNNN 679
 606 LKEGGQDELPSLEKKLLGLIGLGEKPNISVLSESY--KLNS-----TSDSRTINRDR 656
 680 MNS-----INNSNNSNYSYKS-----NYSNOSISDV-QIRYVENMDT 717
 657 VHSFPQPLNNLNPMPFVKSPPKRVVFSKSPKRVAFDSSMSGDSGASAYNSPIQT 716
 718 SKKNDNITFDALSCDNNNYPNITNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 774
 717 SKLKNMNF-----NTMHPSTPAKRPEKNQIDIVE--INTLSPVSPMLLEDXP 763


```

OY 775 ---LYAVNSHLFOPDNNKNTSNIONINTNNKNOGQVNY-----SNMFC 816
DB 764 EPGLLISPLEKKQEVNSESTQDLDLEDSTDSVSPHLDITDLDGSSPVKPKDLDNFS 823
OY 817 HYNLDKNVYLIDLNNEKQKRNHCDDNNIIIONRNDPEKKKTKFNNTNIVIVNNNGN 876
DB 824 RANMDSPEFI--LNNEA-----IHNFDSKPKTRQSLSLTTLHLSNPANILRSLSMAE 876
OY 877 NNSPR 881
DB 877 NEEK 881

RESULT 32
KF11 HUMAN STANDARD; PRT; 1057 AA.
ID AC P52732; Q15716;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kinesin-like protein KIF11 (Kinesin-related motor protein Eg5)
DE (Kinesin-like spindle protein HESP) (thyroid receptor interacting
GN KIF11 OR KNSL1 OR EG5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Butelostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., PHOSPHORYLATION SITE THR-927, AND MUTAGENESIS.
RX MEDLINE=96128120; PubMed=8548803;
RA Blangy A., Lane H.A., D'Herin P., Harper M., Kress M., Nigg E.A.;
RT "Phosphorylation by p34cdc2 regulates spindle association of human
RT Eg5, a kinesin-related motor essential for bipolar spindle formation
RT in vivo."
RL Cell 83:1159-1169 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98369052; PubMed=9701554;
RA Whitehead C.M., Ratner J.B.;
RT "Expanding the role of HsEg5 within the mitotic and post-mitotic
RT phases of the cell cycle."
RL J. Cell Sci. 111:2551-2561 (1998).
RN [3]
RP SEQUENCE OF 819-868 FROM N.A.
RX MEDLINE=95295737; PubMed=7776974;
RA Lee J.W., Choi H.-S., Gyuris J., Brent R., Moore D.D.;
RT "Two classes of proteins dependent on either the presence or absence
RT of thyroid hormone for interaction with the thyroid hormone
RT receptor."
RL Mol. Endocrinol. 9:243-254 (1995).
CC -I- FUNCTION: MOTOR PROTEIN REQUIRED FOR ESTABLISHING A BIPOLAR
CC SPINDLE. BLOCKING OF EG5 PREVENTS CENTROSOME MIGRATION AND ARREST
CC CELLS IN MITOSIS WITH MONOCLONAL MICROTUBULE ARRAYS.
CC -I- SUBUNIT: INTERACTS WITH THE THYROID HORMONE RECEPTOR IN THE
CC PRESENCE OF THYROID HORMONE.
CC -I- PTM: PHOSPHORYLATED EXCLUSIVELY ON SERINE DURING S PHASE, BUT ON
CC BOTH SERINE AND THR-927 DURING MITOSIS, SO CONTROLLING THE
CC ASSOCIATION OF EG5 WITH THE SPINDLE APPARATUS (PROBABLY DURING
CC EARLY PROPHASE).
CC -I- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BINC
CC SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X85137; CAA59449.1; -

```

```

DR EMBL; U37426; AAA86132.1; -
DR EMBL; L40372; AAC41739.1; -
DR PIR; G02157; G02157.
DR PDB; 1IIF; 18-UTL-01.
DR Genew; HGNC:6388; KIF11.
DR GK; P52732; -.
DR MIM; 148760; -.
DR GO; GO:0005871; C:kinesin complex; TAS.
DR GO; GO:0005819; C:spindle; TAS.
DR GO; GO:0004002; F:adenosinetriphosphatase activity; TAS.
DR GO; GO:0003777; F:microtubule motor activity; TAS.
DR GO; GO:0007052; P:mitotic spindle assembly; TAS.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil; Mitosis;
KW Phosphorylation; 3D-structure.
FT DOMAIN 16 363 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 364 480 COILED COIL (POTENTIAL).
FT DOMAIN 737 764 COILED COIL (POTENTIAL).
FT NP_BIND 105 112 ATP (POTENTIAL).
FT MOD_RES 927 927 PHOSPHORYLATION (BY CDC2).
FT MUTAGEN 927 927 T->A: NO MITOTIC PHOSPHORYLATION. NO
FT BINDING TO SPINDLE APPARATUS.
FT CONFLICT 674 676 RNS -> EL (IN REF. 2).
SQ SEQUENCE 1057 AA; 119273 MW; E322F214BEF1601 CRC64;

Query Match 6.5%; Score 447.5; DB 1; Length 1057;
Best Local Similarity 20.9%; Pred. No. 9.1e-13;
Matches 209; Conservative 164; Mismatches 387; Indels 241; Gaps 34;

OY 5 IKVVRKRPSELEKKKKSDITVKNCTLYIDEPYKYD-----MTKTEHEEPIYDK 59
DB 19 IQVVVRCPPEFLERKASAHISIV-----DPVKEKSVRTGGLADRSRKTYTFDM 70
OY 60 VEDDTVNFVYENTIELIIDLKENGCVSCFAYGQSGKTYTMG-----SQPYQSD 115
DB 71 VEGASTQVIDYRVRVCP-ILDEVIMGYNCTIFAYGQSGKTYFMGERSPNEYTWEE 129
OY 116 TP--GIFQYAGDIFFTFLNIYDKNTGIFISFEIYCGKLYDL-----QKRVVAAL 167
DB 130 DPLAGIIPRLTHQFEKLT--DNGTERSVAVSLIETINELFDLINSQVSELRQMFDD 187
OY 168 ENKKEVAVVDKILRVLYTEBELIKMIDGVLLKIGVNSQNDSSSHALINIDK--- 224
DB 188 PRNRKGVILKLEIEITVHNKDEVYQILEKGAARKTTAATLMNAVSSRSVSVTIMKE 247
OY 225 ---DINKNTSLGKIAFTIDLAGSESGADTVSONKQOTDGNINNSLLATKECIRAMSDK 281
DB 248 TTIGBELVAKIGKMLVDLWLAGSENIIGSGADKARERNG--NINOSLTLGRVITAVERT 306
OY 282 NHIFPDSSELTKVLRDIFVGSKSIMTANISPTISCEOTLNTLYSRVNFKNKSTCI 341
DB 307 PHVYRESKLRILQDSIGGTRTSIATISPAIINEETSTLEYVHRAKNIILNK----- 362
OY 342 NEEDDTERTISLDSKSENNASSIENVVVK-----SNHLSNNN 383
DB 363 -----PEVYQKLTAKKALIKETEEIERLKRDLAAREKNGYVISEEN 404
OY 384 NKIRKGIKNDKIER--NNILKNSFQDPREGFSTPKQYSSLNDIDIK---KXKKGLI 438
DB 405 FRVWSGKLTVOEBOIVELIEXIGAVEELNRVTELF--MDVKNLDDQKSLDQKTOELE 462
OY 439 NYKSTLYV-----DNTINKGNNNNNNNNNDNNNDNNNNNNNNSSSMV 483
DB 463 TTQGHLDGTEKQLVKEEITTSALSTEBKLDHDAASKLNTLYVEETTVQVSGHSLDKRKA 522
OY 484 NNMNINHMNINNNIN--VANNNNNNNNNNNNNNNNHNLPGQVYFTD--TSDFSLLDMN 539
DB 523 VDOHNAEQDIFGKLNLSLFFNNMEELIKDGSKOKAMLEVAKTLFGNLLSSVSALDIT 582

```


QY 540 C-----HINNDKSIFLHKKNL-----RDNIKLNCSQINMKKKNL-----579
 DB 583 TVALGSLTSPENSTHVSQFNNMLKEQSLAEKSVLQELIVALTDLSSLEMTLSP 642
 QY 580 -----HARSVSKLTMF-SYDPQKCKNTFKS-----NINMEDTPTKOLLYES 625
 DB 643 TVVSLKINSQLKHIPTSLTADKIDQKKRNSDGLSLCNHLEQENTLCSLESQ 702
 QY 626 RANSNNMGNVLGKNTKTHD-----ISTDENNN-----DNKINNGVI 664
 DB 703 KQCGNLTED--LKTQTHSQELCKMLNLTWTERFCALEKCEKENTQKPLSVQENIQQSK 760
 QY 665 NIINNSVNS-----INNSNM-----NSINNSNNSIYKSNVNSQSI 704
 DB 761 DIVAKMTEHSCFKCADSDGFSQELRNQESTKLVESSVAKSDGLNGLNLEKISQETQRC 820
 QY 705 SDVQIRYV-----NEMD-----TSNKNNDNI 725
 DB 821 ESNLTRFVYFSEQVWSLSNEEQELHMLEVVSQCEASSSDITEKSDGRKAHEKQHNI 880
 QY 726 FPDASCDNNMYPITNNNNNN-----NNNNNNNNIDVENNNNDGTNSMKLY 776
 DB 881 FLDDMTIDEXL--IAQNLBNLETKIGLTKLNCFLBODLKDIP-----GTPPQKSY 933
 QY 777 AYNS-----HNLFPQDNKNTSNINININQNDQNV 809
 DB 934 LYPTLVTRTEPREHLLDQKRPBELMLMNCSENNNEETI 974

RESULT 33
 ID KINH MOUSE STANDARD; PRT; 963 AA.
 AC 061768; 008711; 061580;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Kinesin heavy chain. (ubiquitous kinesin heavy chain) (UKHC).
 GN KTF5B OR KNS1 OR KHCS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 ON NCBI_TaxId=10090;
 RX NCBI_TaxId=10090;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=9726654; PubMed=9112396;
 RA Meng Y.X.; Wilson G.W.; Avery M.C.; Varden C.H.; Balczon R.;
 RT "Suppression of the expression of a pancreatic beta-cell form of the
 RT kinesin heavy chain by antisense oligonucleotides inhibits insulin
 RT secretion from primary cultures of mouse beta-cells.";
 RL Endocrinology 138:1979-1987 (1997).
 RN [2]
 RP SEQUENCE OF 1-861 FROM N.A.
 RC STRAIN=BALB/c;
 RX MEDLINE=94224818; PubMed=8170981;
 RA Gudkov A.V.; Kazakov A.R.; Thimmapaya R.; Axenovich S.A.; Mazo I.A.;
 RA Robinson I.B.;
 RT "Cloning mammalian genes by expression selection of genetic
 RT suppressor elements: association of kinesin with drug resistance and
 RT cell immortalization.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:3744-3748 (1994).
 CC -1- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
 CC PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT. MAY BE
 CC INVOLVED IN THE MECHANISMS OF GROWTH ARREST INDUCED BY EXPOSURE TO
 CC DNA-DAMAGING DRUGS OR BY CELLULAR SENESECE.
 CC -1- SUBUNIT: OLIGOMER COMPOSED OF TWO HEAVY CHAINS AND TWO LIGHT
 CC CHAINS.
 CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS: A LARGE GLOBULAR N-
 CC TERMINAL DOMAIN WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY OF
 CC KINESIN (IT HYDROLYZES ATP AND BINDS MICROTUBULE), A CENTRAL
 CC ALPHA-HELICAL COILED COIL DOMAIN THAT MEDIATES THE HEAVY CHAIN
 CC DIMERIZATION, AND A SMALL GLOBULAR C-TERMINAL DOMAIN WHICH

CC INTERACTS WITH OTHER PROTEINS (SUCH AS THE KINESIN LIGHT CHAINS),
 CC VESICLES AND MEMBRANOUS ORGANELLES.
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U86090; AAB53940.1; -;
 DR EMBL; L27153; AAA20133.1; -;
 DR EMBL; L29223; AAA20132.2; -;
 DR HSSP; P31176; 1BG2.
 DR MCD; MGI:1098268; K15b.
 DR GO; GO:0005515; F:protein binding activity; IPI.
 DR GO; GO:0007028; P:cytoplasm organization and biogenesis; IMP.
 DR InterPro; IPR001752; Kinesin_motor.
 DR Pfam; PF00225; kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
 KW Motor protein; Microtubules; ATP-binding; Coiled coil.
 FT DOMAIN 1 328 KINESIN-MOTOR (BY SIMILARITY).
 FT 329 914
 FT DOMAIN 915 963
 FT NP BIND 85 92
 FT CONFLICT 20 21
 FT CONFLICT 40 40
 FT CONFLICT 75 75
 FT CONFLICT 191 191
 FT CONFLICT 255 255
 FT CONFLICT 396 396
 FT CONFLICT 403 403
 FT CONFLICT 451 451
 FT CONFLICT 520 520
 FT CONFLICT 523 523
 FT CONFLICT 653 653
 FT CONFLICT 664 664
 FT CONFLICT 744 744
 FT CONFLICT 814 814
 FT CONFLICT 865 865
 FT CONFLICT 881 881
 SQ SEQUENCE 963 AA; 109549 MW; 718CDBD102E0475E CRC64;
 Query Match 6.5%; Score 447; DB 1; Length 963;
 Best Local Similarity 21.3%; Pred. No. 8 6e-13;
 Matches 211; Conservative 177; Mismatches 414; Indels 188; Gaps 37;
 QY 5 IKVVRKRLPLSELEKKKSDIITVKNQNTLYIDREPRYKVMYTERHEFIVDKYFDDT 64
 DB 9 IKVRCFRPLNBESEVNRGDK-----YV--AKFGQEDTVVIAKRYAFDRVQSS 55
 QY 65 VDNFTYENTIKRLIIDLYE--NGCVSCFAYQQTSSGTYTTLGS--QPYGSDTPGIF 120
 DB 56 TSOEQVYVNCARKIVDVEGVNGTI---FAYQQTSSGKTHTWEGKLHPDEGGIIPRIY 112
 QY 121 QVAAAGDIFPFLINLYDQNTKGIPISEFYIYCGLYLLQKRVKVALENGKKEV-VVKDL 179
 DB 113 Q-----DIFVITYSMDENLEPHIVSIFELYLKIRLLDVSKTNLVSVEHDKRVPYVKG 168
 QY 180 KILRVLTKEELIKMIDGVLLRKIGVNSQNDSSRSALINTDLKINNTS---LQKIA 236
 DB 169 TERFVCSPEVMDTIDEGKSNRIVAATNNMNSRSHSIFLINVQENQTEQKLGSKY 228
 QY 237 FTLASSKQADTVSQKQOTQDGNINSLAKKCIQAMSDKXHIIPRDSLETKVLR 296
 DB 229 LVDLASSEKVSQKTAGG-AVLDEAKNINSLALGNVIALAEGSTVYPRDSKMTRILO 287


```

Db 823 -----GTTGKPKYKPNSSHDQISGLVLIADHCTIKV--FCGTYSIIPVEAKTYVN 872
OY 544 -----NNDKSI-----FLHKGNLRDNITLKRSSODNIMNK-----KK 577
Db 873 GKILLETTVLRHGDVRLVIGGDHIFRHHPEVQ--GKXPSGHDITISGPPDPEPAK 929
OY 578 NLHARHSVSKLTMSFYDPQ--KKNDTPFKSNINK--MEDNTPKDIYLSERHVMNG 633
Db 930 ELLMAQR--SQLEAEIKEAQULAKEMMGQIQAKEMAQOELSSQAAVESK----- 979
OY 634 NVLLGKNTKTHDITSDENHNKINNGVNIINNSNVNSINNSNMNSINNSNMNSI 693
Db 980 -----IKALEAEIRESQKQKQOELNNQKA--NHKILELEKAKOHLBOETIYVNRKL 1029
OY 694 YKSNYNSNOSISDPQIRYVNMEDTSPKNNKNNINFPDASCNNNPNPTNNNNNNNNNN 753
Db 1030 EMETLATKQLEHDSIRHARILEAELEKOKIAKV----- 1065
OY 754 NNNNIDVENYNNNDGTN-----NSMKL-----YVNSHNLFOPDNMK 790
Db 1066 ---QILQONNNNDKFTVQGTWSSMKLSMWIOEANAISSKLTQYVFGHDSIDKSSSD 1122
OY 791 NTSNIQNT-----NTNKNQDGNVYNSMNFCHYNLNDKNYL 826
Db 1123 TSIRVRNLKLGISTFWSLEKESKLAAMELYESNGSNGEDA--FCDPEDEWEPDI 1177
OY 827 IDL-----NNKEQDKNKHGCDNINONNDPEKKKKTTPYNNNNIYVNNMNGN 877
Db 1178 TDAPVSSLRRRSRLMKNRIRISGCLHDI--QVHPIKULSHSSHGMDKXSTIYSN 1232
OY 878 NSPRMKYGLC-----GS-----HTSIDMKKNEM-----KNEMKDNEM 911
Db 1233 SAESFLPGLIKELIGSLDFPGQSYDEBRTIADSLNSFLKINGLFAISKAHEQDEBS 1292
OY 912 KDNHISNNNNSSSSSSNNNNIYNNINDDTFONDYCHANDNTTIRKKNNTNI--NSNTY 969
Db 1293 QDNLFFSDDRAIOBL-----TIQOT--ACAPEQLVLMKHMWLSDLPCNTNIA 1335
OY 970 Q-NDIIYTNISLNDY-----MSNTL 989
Db 1336 RLEDELRQEVKKYGLQFLQGCCLDISSMIKEAQKNAIQIYQAVKYVQGLAVLKGS 1395
OY 990 LHEKEKYTPYLTSTNEDIYNKEME--GKHRLDD--QDKYDNDNNNNVNNKNNVNN 1044
Db 1396 LHFLEGNKNAASVQEEFMDAVCDGVLGKMKIILDSGLEKAKELQHELPQCTKAEVTE 1455
OY 1045 VDNNDVND--NNVNDNDKNNVNDVNDVNDVNDVNDVNDVNDVNDVNDVNDVND 1097
Db 1456 MKNAMGLIRSLNIRLFAESKIKSFRRQV--QENPFYODFKRNVNAPFELK----- 1505
OY 1098 IINNCCNSLDISSMYDTEKILANNILSKYK--AEKONVIAKKYINEDIKNSLE--EL 1151
Db 1506 -LKHCL-----EKALEIITISALKGCHSDINLQTCVESIRNLASDFYSDFSV 1551
OY 1152 DKTAOSIYERKAVLTKLLLLFKKNVDTQINNETSLDKDLVWCHICNNPD 1203
Db 1552 PSTSVGSYSER-----VTHIVHQLBSLAKSLRCLFSESESPD 1589

```

```

RESULT 35
KF3A MOUSE
ID KF3A MOUSE STANDARD; PRT; 701 AA.
AC P28741;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Kinesin-like protein KIF3A (Microtubule plus end-directed kinesin motor 3A).
GN KIF3A OR KIF3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93077686; PubMed=1447303;
RA Mizawa H., Sekine Y., Takemura R., Zhang Z., Nangaku M., Hirokawa N.;
RT "Kinesin family in murine central nervous system."
RL J. Cell Biol. 119:1287-1296(1992).
CC -1- FUNCTION: MICROTUBULE-BASED ANTEROGRADE TRANSLATOR FOR MEMBRANOUS ORANELLES. PLUS END-DIRECTED MICROTUBULE SLIDING ACTIVITY IN VITRO.
CC -1- SUBUNIT: HETERODIMER OF KIF3A AND KIF3B.
CC -1- TISSUE SPECIFICITY: EXPRESSED ALMOST EXCLUSIVELY IN ADULT BRAIN TISSUE (MAINLY IN THE CEREBELLAR GRANULAR LAYER) WITHIN A SINGLE TYPE OF NEURONAL CELL.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN II SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb.ch).
CC -----
DR EMBL; D12645; BAA02166.1; -
DR PIR; B44259; B44259.
DR HSSP; P17119; 3KAR.
DR WGD; WGI:107689; Kif3a.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC. 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
DR Motor protein; Microtubules; ATP-binding; Coiled coil; Neurone.
KW DOMAIN 1 350 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 351 586 COILED COIL (BY SIMILARITY).
FT NP_BIND 587 701 GLOBULAR.
FT NP_BIND 100 107 ATP (BY SIMILARITY).
FT DOMAIN 442 445 POLY-GLU.
FT DOMAIN 509 512 POLY-ARG.
SQ SEQUENCE 701 AA; 80167 MW; 2405872DP2D85A29 CRC64;

```

```

Query Match 6.5%; Score 445.5; DB 1; Length 701;
Best Local Similarity 29.6%; Pred. No. 7.1e-13;
Matches 145; Conservative 84; Mismatches 196; Indels 65; Gaps 16;

```

```

OY 5 IKVVRKRPLELEKKKQSDIITVKNNTCTLYIDEPKYVDWTKYIERH-----FIVDKV 60
Db 15 VAVVVRKRLNREKSMCRQAVSV-----DEMKGITTHAKDSSNEPKTFTFDV 66
OY 61 FDDTVNFTVYENTIKPLIIDYE--NGCVSCFAYGQSGKTYMTLGSQPYGSDPFG 118
Db 67 FEPESKQDLYVNLTPAPIIDSVLEGNGTI--FAYGQGTGKFTTWGVR--AVGGLRG 121
OY 119 IFQYAAGDIFTPLNIYDQNTKGIIFISPEYICGLYDLQK--RKMAALENGKEVYV 176
Db 122 VIPNSFAHIFGHIAAEGDTPFVAVSYLAIYNEBVRDLGDOQTORLEVKRPDVGVI 181
OY 177 KDLKILRVLTKEELIKMIDGVLTKRIGVNSONDSRSSHAATIDLK-----DINKVT 230
Db 182 KQLSAVYVNMADMDRINTLGHKRSVGTATNNHSSSHALFTTIESSEGVQGNHVV 241
OY 231 SLGKIAFIDLAGSEKADTVSONKQOTDGANINSLALKECIKRAM--DSDKNHLPFDS 289
Db 242 RMGKLHLVDLAGSEKQATGATGRK--EATKINISLTLGNVISALVDGKSTHVPYNS 300
OY 290 ELTKLRLDIPVGSKSMIANISPTISCEQOTLNTKRSRYKRNKSTCINED----- 345
Db 301 KTLRLQSLGSKTMKCANIGPADYVDITSLRYANRKNKIKNKR--INEDPKDAL 359

```



```

QY 346 ----DTNTERISILDSKSEMNASSIENVVIKSNHLLNNNNKINKINDKIERNNIL 401
DB 360 LROFOKEIEELKKLEBEEVSGSDISGEDEDEGEIGDEK--RKRRDQAGKKV 417
QY 402 KNR-----SPDKREGFTSTFGKYSSLANDIKIKNK-----KKGLI---NYKST 443
DB 418 PDKVEMQAKIDBERKXLETYL-----DWEERERNKARAELEBRERKDLKAOQHOSL 470
QY 444 LYNDNTINKK 453
DB 471 LEKLALAEKK 480

RESULT 36
KFSC_MOUSE STANDARD: PRT: 956 AA.
ID_KFSC_MOUSE
AC P28738; Q922F8;
AT 01-DEC-1992 (Rel. 24, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE kinesin heavy chain isoform 5C (kinesin heavy chain neuron-specific
  2).
GN KIF5C OR NKHC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclerogonathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP PRELIMINARY SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Brain;
RA Kato K.;
RT "A collection of cDNA clones with specific expression patterns in
  mouse brain."
RL Eur. J. Neurosci. 2:704-711(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99000842; PubMed=9782088;
RA Xia C., Rahman A., Yang Z., Goldstein L.S.B.;
RT "Chromosomal localization reveals three kinesin heavy chain genes in
  mouse."
RL Genomics 52:209-213(1998).
CC -I- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
  PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT.
CC -I- SUBUNIT: OLIGOMER COMPOSED OF TWO HEAVY CHAINS AND TWO LIGHT
  CHAINS.
CC -I- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS: A LARGE GLOBULAR N-
  TERMINAL DOMAIN WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY OF
  KINESIN (IT HYDROLYSES ATP AND BINDS MICROTUBULE), A CENTRAL
  ALPHA-HELICAL COILED COIL DOMAIN THAT MEDIATES THE HEAVY CHAIN
  DIMERIZATION, AND A SMALL GLOBULAR C-TERMINAL DOMAIN WHICH
  INTERACTS WITH OTHER PROTEINS (SUCH AS THE KINESIN LIGHT CHAINS),
  VESICLES AND MEMBRANOUS ORGANELLES.
CC -I- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
  SUBFAMILY.
CC -I- CAUTION: REF.1 SEQUENCE SEEMS TO HAVE BEEN A HYBRID OF A KIF5A
  AND A KIF5C SEQUENCE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation
  at the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X61435; CAA43677.1; -.
DR EMBL, AF067180; AAC79804.1; -.
DR HSPF, P56536; 2XIN.
DR MGI, MGI:1098269; Kif5c.
DR GO, GO:0005515; F:protein binding activity; IPT.
DR GO, GO:0008045; P:motor axon guidance; IMP.
DR InterPro, IPR001752; kinesin_motor.

```

```

DR Pfam, PF00225; kinesin, 1.
DR PRINTS, PR00380; KINESINHEAVY.
DR SMART, SM00129; KISC: 1.
DR PROSITE, PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE, PS50067; KINESIN MOTOR DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
FT DOMAIN 1 386
FT DOMAIN 406 923
FT DOMAIN 859 956
FT DOMAIN 174 315
FT NP BIND 86 93
FT BIND 956 AA; 109240 MW; D5A8C701A2911AB2 CRC64;
SQ SEQUENCE

Query Match 6.5%; Score 443.5; DB 1; Length 956;
Best Local Similarity 21.8%; Pred. No. 1,2e-12;
Matches 216; Conservative 162; Mismatches 426; Indels 185; Gaps 38;

QY 5 IKVVRRRRPLSELEKKKSDIITVKNCTLYIDPRRYKDMTKYI-ERHEFLVDKVPDD 63
DB 9 IKVMCRPRPLNEAELIRGDKFI-----DKFKGEETVVIQGGKPYVFDHYLPP 55
QY 64 TVDNFYENTITPLIIDLYE--NGCVSCFAYGSGGTMTLSQPYGSDTGIFQ 121
DB 56 NTTQEOVYNACAIQYVDVLEGYNGL--FAVGQTSSEKTHMEGK--LHPQLMGILP 110
QY 122 YAAADITPTINIDYDKNTKGIFFISFYEYCGKLYDLQKKRVAALENGKKEV-VVXDLK 180
DB 111 RIADIDIDHYSMDENIEFHIXSYPEIYIDKIRDLIDVSKTULVAHEDKGRVPYVAGCT 170
QY 181 ILRLTYEELLKIMIDVLLRKIGVNSQNDSSRSHAILNIDLKINKNTS---LGIAP 237
DB 171 ERFVSPSEVWDVDEBKARHVAVTMMNHSRSHSIFLINIKOEVEFEKLSGLLYL 230
QY 238 IDLAGSRGADYTSQNKQOTDGANINRSLALKECIRAM-DDDKNHIPRDSELTFKVR 296
DB 231 VDLAGESEKVKRTAGEG-AVLDEAKINKSLSALGNVTSALAEGTKTHVPRDSKTRILQ 289
QY 297 DIFVGSKSLIMINISPTISCCBOTLNTLYSSRVKFNKSKTCIN-----E 343
DB 290 DSLQGNKRTTIVICSSVSNEAETKSTLFGQRAKIKR-TYVNLIELAEWKKKYK 348
QY 344 EDDTNTERTISILDSKSEMN-----ASSIENVVIKSNHLLNNNNK1--NEKINDX 394
DB 349 EKEGNKALKSVLQHLMEINRWENGEAVPEDEQISAKDHSLKPCDPTPIIDNTPVVDG 408
QY 395 IERNILKANSFQKPRREGFTSTF---GYXSLNDIDIKIKKKKKGLINXSTLYNDTI 450
DB 409 ISAEK-----EKYDEITSLYKQDLDDKODEINQOSQLAEKLAQQLDODELASTRRD 461
QY 451 NKKNNNNNNNNNDNNNDNNNNNNNN-----NNDSSSMVNN--MINMTNNNNINN 498
DB 462 YEKIQBELTRQIENEAKEVEKVALAEELAVNYQKQOEVDKTRAEVQLDELAQK 521
QY 499 INVNNNNNNNNNNNNNNHNLPPD---NYAFPD-----TSPSSLDMMNCHLNN 544
DB 522 TTTLTTRQELSQOELSNHOKKATEILNLLKDLIEGIIIGTINVTKLADNGVIEE 581
QY 545 NDKSIFLAKKLRNLIKRS-----SCNINMKKKNLHLARHSGSKLTMPSPDOK 599
DB 582 EFTMARLYISPKKEVSLVRSKQLBSAQMDSRRKN--ASRETLAQCQLLSQHEAK 638
QY 600 NKDNTFPKSN1-----NKMEDNTPRDILYESHNVSNNGVNLGLKNTTHHDISTKDEHN 655
DB 639 IKSITDYNQNEQKRQLQEE--QDSL--SEELAKLA-----QKMEHVSQD----- 683
QY 656 DNKINGVINIINNSVNSINNSNNSINNSNNSINNSNNSIYKSNVNSQGISDVGIRYVN-- 713
DB 684 --KEKEHLTRQDAEYVKALEQOMESHREAHQQLRLDEIEKORIID-ETRDINQK 740
QY 714 -EMDTSKNDNITFPDAISCDNNMYPNTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 772
DB 741 LQLBERLSSD-----YKLTIEDEREVVKLEK 769

```


QY 773 MKLYAVNSHMLFQPDNNKNTSNIONINTKNNKQDGNVYNSMFCYVNLNDKNYIIDLNNK 832
 DB 770 LLL-----NDKR-----EQAREDLKGLBEFVYSIELQTLH-NLR-KLFPQDLTTR 811
 QY 833 EQDKXNHHGDDNNIIOKRNDEKFKKTNFYNN--NIVYNNNNNNNS-----PKMY 884
 DB 812 VKKSVELDSDG-----GSAAGQOKSIFENNLEOLTKYKQLVRONADRLCELPLEK 866
 QY 885 GLGSHTSIDNNKNNKNNKNNKNNKMD 913
 DB 867 RL---RATAEKRVKALLESALKKAKENAMRD 892

RESULT 37
 ID_KF4A_XENLA STANDARD; PRT; 1226 AA.
 AC 091784; 09PST0; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chromosome-associated kinesin KLP1 (Chromokinesin).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
 RE SPECIFICITY.
 RC TISSUE=Oocyte;
 RX MEDLINE=95236444; PubMed=7720067;
 RA Vernos I., Raats J., Hiran T., Heasman J., Karsenti E., Wylie C.;
 RT "Xklp1, a chromosomal kinesin-like protein essential for
 spindle organization and chromosome positioning.";
 RL Cell 81:117-127(1995).
 RN [2]
 RP SEQUENCE OF 9-338 FROM N.A.
 RX MEDLINE=93246065; PubMed=8482413;
 RA Vernos I., Heasman J., Wylie C.;
 RT "Multiple kinesin-like transcripts in Xenopus oocytes.";
 RL Dev. Biol. 157:232-239(1993).
 CC -1- FUNCTION: REQUIRED FOR MITOTIC CHROMOSOMAL POSITIONING AND BIPOLAR
 SPINDLE STABILIZATION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR, ASSOCIATED WITH MITOTIC
 CHROMOSOMES.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN OOCYTES, EGGS, TESTES AND
 BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
 CC CHROMOKINESIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DB EMBL; X82012; CAA57539.1; -.
 DB PIR; I51617; I51617.
 DB HSSP; P17119; 3KAR.
 DB InterPro; IPR001752; kinesin_motor.
 DB Pfam; PF00225; kinesin_1.
 DB PRINTS; PR00380; KINESINHEAVY.
 DB SMART; SM00129; KISC; 1.
 DB PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 DB PROSITE; PS00607; KINESIN MOTOR DOMAIN2; 1.
 KM Motor protein; Microtubules; ATP-binding; DNA-binding;
 KM Nuclear protein; Coiled coil.
 FT DOMAIN 1 350 KINESIN-MOTOR
 FT DOMAIN 351 1006 COILED COIL (BY SIMILARITY).
 FT DOMAIN 1007 1226 GLOBULAR.

FT NP BIND 87 94 ATP (POTENTIAL).
 FT CONFLICT 163 163 I -> L (IN REF. 2).
 SQ SEQUENCE 1226 AA; 138923 MM; 7F0275FCF3316697 CRC64;
 Query Match 6.5%; Score 442.5; DB 1; Length 1226;
 Best Local Similarity 24.9%; Pred. No. 1,8e-12;
 Matches 174; Conservative 118; Mismatches 275; Indels 131; Gaps 24;

QY 5 IKVVRKRPLSELEKKKXKSDITTVKNN-----CTIYDEPRKVMYTKIEHFEIVDK 59
 DB 9 VRALACRPL-----VPKENNECKKCLFVFEQEQVITYGTE---KSTFYDY 52
 QY 60 VFDDTVNDFVTYENTIKPLIIDLYENGCVCSFPAVQTSQSGKTYTLMGSPQYQSDPTGI 119
 DB 53 VFDPASAEQEEVYNSAVALPKGLFK-GYNATVALAVQTSQSGKTYSGMGAATHQNEPVT 111
 QY 120 FQYAAAGDIFPLNIYQDKNTK-GIFSPFEIYQCKLYDL---QKRVVALENGKEV 174
 DB 112 GVIPRTVIALFREIHRPEWEMFLKYSYLEIYNEEILDLIYARDKNTISIREDPREGI 171
 QY 175 VVVDLKLRLVLTKEELILKMIIDGVLLRKIGVNSQNDSSRSNAILNIDL---KDINKNTS 231
 DB 172 KIGGLTERDYKTALDITLSLBOGNSRVTASTMNSQSSRSNAILFTISIQKKEGDKNS 231
 QY 232 L-GKIAFIDLAGSERGADTVSQNKQOTQDQANINRSLALKECIRAM-DSDK--NHI PFR 287
 DB 232 FRSKLHLVDLAGSERQKTKRABGDRLK-EGISINRGLLIGNVTSALGDESKKGFVPR 290
 QY 288 DSELTKVLADIFPKGKSLMINIANSPTISCCEPTLNTIAYSSVVKPKKSTINEEDT 347
 DB 291 DSKLTRLDPSLQSGNSHTTLMACVSPADSNMEETLTLRYADARKIKKPIVNTDPOAA 350
 QY 348 NTERIS-----ILDSGSEMANSSIENVYTKSNHLSNNNNKINRGKINDKIERN 398
 DB 351 ELQRLKQVQELVLLQAHGTL---PVLNSPESENLOSIMERK-NLEKNGKLSRE 406
 QY 399 NILKNSFKPREGFTSTFGKYSLSNDIDKI---KKNKKGLINYKSTLYNDNTINKK 454
 DB 407 -LGEAAVQ-----TAQF-----LEKIINTEQNEELGSKMELQHAACKYVNLQ 450
 QY 455 NNNNNNNNDNNNDNNNDNNNNNNN-----NDSSKVNMMIYHMINNNINNNNNNNNN 509
 DB 451 LVETLEDOELKQWVEVIQNIQVTVQLQDESSGIAQSIAMDEEASFVPEDESGEKS 510
 QY 510 NNNNNNNHNLPOPNVAFDTSDSSSLDMNCH-----NNNDKSIPLHKKRLRN 559
 DB 511 SDGFTTHALRQQLS-----KELIELNKALVKEALAKKQAQNDRLQEP1QSEYLYNN 563
 QY 560 ILLKNNSSCDNINNNKKNMILAHRSVSGSKLITWFSYDPOKNKNDTPFK----- 607
 DB 564 I-KHLESEVGLQKEKELELLALHSA-----KDDNQAKLSERRRKLQEL 608
 QY 608 -----SNINMEDNTPDIIYESRNSNM 631
 DB 609 EGQMTLKKKLGEOGSKLKLRESTEKTAVAKMQEIQGM 646

RESULT 38
 ID_EG52_XENLA STANDARD; PRT; 1067 AA.
 AC 091783;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Kinesin-related motor protein Eg5 2.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eg;

RX MEDLINE=94299012; PubMed=8026619;
 RA Holliston E., le Guellec R., Kress M., Philippe M., le Guellec K.;
 RT "The kinesin-related protein Egs associates with both interphase and
 spindle microtubules during Xenopus early development.";
 RL Dev. Biol. 164:147-159 (1994).
 CC -1- FUNCTION: PLUS END-DIRECTED MOTOR PROTEIN REQUIRED FOR
 ESTABLISHING A BIPOLEAR SPINDLE. ASSOCIATES WITH BOTH INTERPHASE
 AND SPINDLE MICROTUBULES. MAY BE INVOLVED IN NUCLEAR DIVISIONS
 TAKING PLACE DURING THE DEVELOPMENT OF UNFERTILIZED EGGS.
 CC -1- SUBUNIT: HETEROTETRAMER OF TWO HEAVY AND TWO LIGHT CHAINS (BY
 SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CONCENTRATED AROUND THE POLAR ENDS OF
 BOTH MEIOTIC AND MITOTIC SPINDLES.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN UNFERTILIZED EGGS,
 ESPECIALLY IN THE GERMINAL VESICLE AND IN THE RADIAL YOLK-POOR
 CHANNELS. ALSO PRESENT IN TESTIS.
 CC -1- DEVELOPMENTAL STAGE: ACCUMULATES DURING THE LATTER STAGES OF
 OOCYTESIS AND LEVELS INCREASE THREE-FOLD DURING OOCYTE MATURATION.
 CC LEVELS DECREASE AFTER FERTILIZATION.
 CC -1- PMT: PHOSPHORYLATION DURING MITOSIS AT THR-937 CONTROLS THE
 ASSOCIATION OF EGS WITH THE SPINDLE APPARATUS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIOC
 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement and this agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib.ch).
 CC -----
 DR EMBL, X71864; CAAS0695.1; -.
 DR PIR, I51616; S33417.
 DR HSSP, P17119; 3KAR.
 DR InterPro: IPR001752; kinesin_motor.
 DR Pfam: PF00225; kinesin_1.
 DR PRINTS: PR00380; KINESINHEAVY.
 DR SMART, SM00129; KISC.1.
 DR PROSITE, PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE, PS0067; KINESIN MOTOR DOMAIN2; 1.
 DR Motor protein; Microtubules; ATP-binding; Coiled coil; Mitosis;
 KW Phosphorylation.
 FT DOMAIN 16 363 KINESIN-MOTOR (BY SIMILARITY).
 FT DOMAIN 364 487 COILED COIL (POTENTIAL).
 FT DOMAIN 692 736 COILED COIL (POTENTIAL).
 FT DOMAIN 884 920 COILED COIL (POTENTIAL).
 FT DOMAIN 1040 1067 COILED COIL (POTENTIAL).
 FT NP BIND 105 112 ATP (BY SIMILARITY).
 FT MOD_RES 937 937 PHOSPHORYLATION (BY CDC2) (BY
 SIMILARITY).
 SQ SEQUENCE 1067 AA; 119885 MW; 48AA47DBE2559115 CRC64; .
 Query Match 6.4%; Score 439.5; DB 1; Length 1067;
 Best Local Similarity 22.1%; Pred. No. 2e-12;
 Matches 198; Conservative 156; Mismatches 327; Indels 213; Gaps 31;

OY 217 AINIDLK-----DINKNTSLGKIAFIIDAGSERGADTVSONKQOTQTDGAINIRSLAL 270
 DB 237 SVSVTHIMMETTIDGSELVYKIGKUNVDLAGSENGRSGAVDKAREAG-NINQSLTLT 295
 OY 271 KECIRAMDSPKNHIFPDSSELTVALDI FVGSKSIIMIANISPTISCEOTLNTLRYSR 330
 DB 296 GRVITLALVERAPHIPYRESKLTIRLOPSLGRTTSIIATVSPASINLEETMSTLDYASR 355
 OY 331 VKPKNKSTGICNEEDPTNTERISILD-SKSENMASISENVVAKSNHLNNNNKRNKG 389
 DB 356 AKNIMNKP-----EVNQKLTAKALIKETIEIERKELATAREKNGVYLSNENEQ-OG 410
 OY 390 KI-----NDKIER-----NNILK 402
 DB 411 KVLSDGEMITEYSKIAMEEIKRIGELFADNKKKELECTILOQCKEKELEATQNNLOE 470
 OY 403 NKSPDKREGFTSTFGYSSLLNDID-KIKKKKKGLINYSTLYNDNTINK-----453
 DB 471 SKE-QLAQEAFV-----VSAMEETTEKHLGTANTLSTVETTRDVSGLHEKLDKRAVE 524
 OY 454 -HNNNNNNNDNNNNNN-----DNNNNNNND-----478
 DB 525 QHNSQVHEHNPBQINRPFSVYIQOTVDEYSVKQGMDFYNSIDDLGASSALSATATA 584
 OY 479 -SSSMVNNMINMINNN 510
 DB 585 VAKSFPAVOETVSGVSHSVBEILKQETLSQAKDDQLKMTAHTGLEQALRTDLLPV 644
 OY 511 -NNNSHNHLLPQNYAFTDTSDFSLDNDNCHLNNDKSLFLAKNLPDNLKORS 566
 DB 645 TAVLDLNSHLSHCQSQSLGVADKID-SHKEDMNSFTEHSRS-LHLRLDSSSL--S 698
 OY 567 SCDNINMKKKNNLHLAR--HSVSKLTMFSDPOKN-----KDNTEFKNIKMEENT 617
 DB 699 SIQSEYSLKEEITATQSTHSEGVNNLISLQNLNLAMETTRQSGFSLSGKQLOESV 758
 OY 618 PKDILYESSRVNNGNVLGLNKNTHHDISTXDENHNDKKNNGVNIINNSVNSINN 677
 DB 759 --GGLQDDLDVSSDAIECI--SSHHSKTEGSAVTEIRQ-----LAGSNWSTLEE 807
 OY 678 SN-----MNSINNSNNNSIYSVNSNOSISDVQIRYVN---EMDSNKN 721
 DB 808 SSKQCEKLTNSINITICQSQQWCSAGQKMDSLBEQVCYLHSSKKQIQTLAKD 861
 RESULT 39
 KFSC_HUMAN STANDARD; PRT; 957 AA.
 ID KFSC_HUMAN
 AC 060282; 095079;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Kinesin heavy chain isoform 5C (Kinesin heavy chain neuron-specific
 2).
 GN KIF5C OR NKHC2 OR KIA0531.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 5:31-39(1998).
 RN [2]
 RP SEQUENCE OF 355-585 FROM N.A.
 RA Engelder S., Sharp A.H., Colomer V., Tokito M.K., Lanahan A.,
 RA Worley P., Holzbaur E.L.F., Ross C.A.;


```

RT "Huntingtin associated protein 1 (HAP1) interacts with the p150glued
subunit of dyactin."
Rl Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
CC PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT.
CC -I- SUBUNIT: OLIGOMER COMPOSED OF TWO HEAVY CHAINS AND TWO LIGHT
CC CHAINS.
CC -I- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN BRAIN, PROSTATE, AND
CC TESTIS, AND MODERATE EXPRESSION IN KIDNEY, SMALL INTESTINE, AND
CC OVARY.
CC -I- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS: A LARGE GLOBULAR N-
CC TERMINAL DOMAIN WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY OF
CC KINESIN (IT HYDROLYZES ATP AND BINDS MICROTUBULE), A CENTRAL
CC ALPHA-HELICAL COILED COIL DOMAIN THAT MEDIATES THE HEAVY CHAIN
CC DIMERIZATION, AND A SMALL GLOBULAR C-TERMINAL DOMAIN WHICH
CC INTERACTS WITH OTHER PROTEINS (SUCH AS THE KINESIN LIGHT CHAINS),
CC VESICLES AND MEMBRANOUS ORGANELLES.
CC -I- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB011103; BAA25457.1; -
DR EMBL; AF010146; AAD01436.1; -
DR HSSP; P56536; 2KIN.
DR Genew; HGNC:6325; KIF5C.
DR MIM; 604593; -
DR GO; GO:0005871; C:kinesin complex; TAS.
DR GO; GO:0003777; F:microtubule motor activity; TAS.
DR GO; GO:0006996; P:organelle organization and biogenesis; TAS.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
FT DOMAIN 1 386 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 406 923 COILED COIL.
FT DOMAIN 859 956 GLOBULAR.
FT NP BIND 86 93 MICROTUBULE-BINDING.
FT NP BIND 355 360 ATP (BY SIMILARITY).
FT CONFLICT 583 585 TLKVI -> STHSAV (IN REF. 2).
FT CONFLICT 583 585 EFT -> DRV (IN REF. 2).
SQ SEQUENCE 957 AA; 109494 MW; A9F25BBLCG94322A CRC64;
Query Match 6.4%; Score 439; DB 1; Length 957;
Best Local Similarity 21.4%; Pred. No. 1.9e-12;
Matches 216; Conservative 173; Mismatches 395; Indels 226; Gaps 41;

```

```

Db 231 VDLAGESEKVTAGAEQ-AVLDEAKNINIKSLSALGNVISALAEKGTIVPYRDSKMTRLQ 289
Qy 297 DIFVGSKSKIMINISPTISCCQTLNTLAYSRRVKNFK- 343
Db 290 DSLGNCRTTIVICSPSVENAEATSTLMFGRAKTKVTSVNLETAEMWKYKE 349
Qy 344 EDDTNT- 371
Db 350 KKKKTKAVNYQHLWELEKRWGRGEAVPEBQISADQKLEPCDTPITDITAPVAGI 409
Qy 372 - 412
Db 410 STEEKEKYDEISLVRQLDDKDEINQSQLAEKQLQWLDQDELLASTRYEKLQEE 469
Qy 413 FTS-TRGKYSLMDIDIKKQKKGGLINT--KSTLYNDNTINKKNNNNNNNDNDNN 469
Db 470 LTRQLIENEAKQDEVEVLQALEBLAVNDQKQVEDEKT- 523
Qy 470 DNNNNNNNDSSMNNNNINM- 523
Db 524 TLTTLTTORELSQI-QELSNHQKGRATEILMLKDLGEIG- 562
Qy 524 YAFETDSFSSLDPMNCHLNNDKSIPLHKQLRDNK-LKNS- 578
Db 563 -GIIGTNDVKTLDVGVIEEFTMARLYISKKSEKSLVNSKQLESQMSNRKN- 620
Qy 579 LHLARISVGSKLTMFSDYPOKNDTFPKSNT--NKMEDNTPKDIYESRVSNMNGN 634
Db 621 --ASERELACQLLISOHEAKIKSLTDYMNQKQKQOLEES--QDSL-SEELAKLRA- 673
Qy 635 VLLGLNKNTHTDSTDEHNNDKINGVNIINNSVNSINNSNNSINNSNSIY 694
Db 674 - 721
Qy 695 KSNYNSNOSISDVQIRYVN---EMDTSNKNNDNIFEDAISCDNMYPNITNNNNNNNN 751
Db 722 RDEIEEKQKIID-EIRDNQKQLQEOEKLSDD- 752
Qy 752 NNNNNNIDVENYNNRGCTNSMCLYAVNSHNLFPQDNNKTSINQINTYKNNQDGVNY 811
Db 753 - 793
Qy 812 SNAFCYNNLNDKYYLLDLNNKEQDKNKHGCDNNIINRNDPEKKKKTNYNNN-NIVI 869
Db 794 ELQTLH-NLR-KLFVDLITRVKSVLELDNDGG- 846
Qy 870 VNNNMGNNS- 913
Db 847 VHQQLVRDADNADLCELPKLEKRL--RATYERVKALESLKKEKENAMD 953

RESULT 40
KININ DROME
ID KININ DROME STANDARD; PRT; 975 AA.
AC P17210; O9V719;
DT 01-AUG-1980 (Rel. 15, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Kinesin heavy chain.
GN KHC OR KIN OR CG7765.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=89168428; PubMed=2522352;
RA Yang J.T., Laymon R.A., Goldstein L.S.B.;
RT "A three-domain structure of kinesin heavy chain revealed by DNA
sequence and microtubule binding analyses."
Cell 56:879-889 (1989).

```


[2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaratunga C.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle B., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 RA Bokora D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busen D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Daventport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Jaisi M., Kalush F., Karpen G.H., Ke Z., Kemsion J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
 RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP MUTAGENESIS.
 RX MEDLINE=93030741; PubMed=1384131;
 RA Gho M., McDonald K., Ganetzky B., Saxton W.M.;
 RT "Effects of kinesin mutations on neuronal functions.";
 RL Science 258:313-316(1992).
 CC -1- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
 CC PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT.
 CC -1- SUBUNIT: OLIGOMER COMPOSED OF TWO HEAVY CHAINS AND TWO LIGHT
 CC CHAINS.
 CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS: A LARGE GLOBULAR N-
 CC TERMINAL DOMAIN WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY OF
 CC KINESIN (IT HYDROLYZES ATP AND BINDS MICROTUBULE), A CENTRAL
 CC ALPHA-HELICAL COILED COIL DOMAIN THAT MEDIATES THE HEAVY CHAIN
 CC DIMERIZATION, AND A SMALL GLOBULAR C-TERMINAL DOMAIN WHICH
 CC INTERACTS WITH OTHER PROTEINS (SUCH AS THE KINESIN LIGHT CHAINS),
 CC VEICLES AND MEMBRANOUS ORGANELLES.
 CC -1- MISCELLANEOUS: MUTANT FLIES DISPLAY IMPAIRED ACTION POTENTIAL
 CC PROPAGATION AND NEUROTRANSMITTER RELEASE AT NEUROMUSCULAR
 CC JUNCTIONS. BUT ARE STILL CAPABLE OF TRANSPORTING CERTAIN
 CC MEMBRANES, INCLUDING SYNAPTIC VESICLES, TO THE NERVE TERMINAL.
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: M24441; AAA28652.1; -
 DR EMBL: AF003807; AAF58029.1; -
 DR PIR: A31497; A31497.
 DR HSSP: P33176; 18G2.
 DR FlyBase: FBgn0001308; Khc.
 DR GO: GO:0005871; C:kinesin complex; IDA.
 DR GO: GO:0008017; F:microtubule binding activity; IDA.
 DR GO: GO:0003774; F:motor activity; NAS.
 DR GO: GO:0007317; P:regulation of pole plasm oskar mRNA localiz. . . ; IMP.
 DR InterPro: IPR001752; Kinesin motor.
 DR PRINTS: PF00225; kinesin. 1.
 DR PRINTS: PR00380; KINESINHEAVY.
 DR PROSITE: PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE: PS00667; KINESIN MOTOR DOMAIN2; 1.
 KM Motor protein; Microtubules; ATP-binding; Coiled coil.
 FT DOMAIN 1 334 KINESIN-MOTOR (BT SIMILARITY).
 FT FT 335 931 COILED COIL.
 FT DOMAIN 932 975 GLOBULAR.
 FT DOMAIN 180 321 MICROTUBULE-BINDING.
 FT NP BIND 92 99 ATP (BY SIMILARITY).
 FT CONFLICT 515 515 A -> T (IN REF. 1).
 SQ SEQUENCE 975 AA; 110399 MW; 24840EF414790888 CRC64;
 Query Match 6.4%; Score 439; DB 1; Length 975;
 Best Local Similarity 21.9%; Pred. No. 1.9e-12;
 Matches 219; Conservative 164; Mismatches 429; Indels 186; Gaps 37;
 5 IKVVRKRPSELEKKKSDITVKNKNTLYIDPRKYDMTYIRHFIYDKVDDT 64
 13 IKVCRFRPLNDSEKSKGKRVVFPNN---VEENCISLAGKY-----LFQKVPKN 62
 65 VDNFTVYENTIKPLIDLYE--NGCVSCFAYGQSGKTYTMGSGPDPIFGY 122
 63 ASQKRVNKAASIVTVLACYNNTI---RAYGTSKTHTMG--VIGDSVAGIIR 117
 123 AAGDIFPLNYDKNTKGIFISFYCYGKLYDLQKRKVAALENGKKEV-VKDKI 181
 118 IVNDIFPHIYMEVNLFEHIVKSYEILYMDKIRDLVSKNLSVHEDKKNVPYKATE 177
 182 LRVLTKEELIKMIDGVLLKRTGVNSQNDSSRHALLNDKQIN---KNTSLGIAF 237
 178 RFVSSPDVPEVIEEGKSNRIIAVTNNHSSRSVFLINVQENLENQKLS-GKLYL 236
 238 IDLAGSRGADTVSONNOTDGA-NINRSLATKECIRAM-DSKNIHPRDSEITKVL 295
 237 VDLGSEKSVKTAGEG--TVLDEKNTKISLALGNVISALADGKTHIPRDSKLRIL 294
 296 RDIIVGSKSIVIMANISPTISCCBOTNTLYRSRVKNFKKSTCINEEDPTNTERISIL 355
 295 QESLGNAARTIVICGSPASNESETSTLDFGRRAKTVKN-VVCNVE--LTAEBWKR 351
 356 DSKSEWNA--SSEIENVVYKSNHLNNSNNKINRGKINDIKERN----- 398
 352 YEKEKKNARLKGVEKLEIELARWGETVKAEOINMEIMEAASPNLEVEAAQDAA 411
 399 -----NLIKKSPEKPREGFTSTGKYSINDIDIKKNGKKGILNYKSTLYN 446
 412 EALAAQATLANASASAVAEQARLATCEGRLYQDLDDKE-EINQOS--QYABOLKE 467
 447 DNTINKKANN 506
 468 QVMOEBELIARAREYETLSEMARIQEENSAEVEYVLALEELAVNYD-QKQSD 526
 507 NNNNNNNSHNHLPOPNVAFPTDS-DFSSLDPNCHLANNDKSLFLKKNLRD----- 558
 527 NKNNDIALNELLOOKQSVFNAASTELQOLKDMSSHOKRITTEMLTNL--LRDGEVGA 584
 559 -----NLIKKNRSCONIKKKNNNLIHLASVSKLTMVSYDPQKNDTPFKSNINK 612
 585 IAPGESSIDLKMSLAGTDAKVEEDFTMALFLSKKTEAKNTAQRCSNVEIQOADSNC 644
 613 MEDNTPKDIILYESRNVNNGNVLLGLNKNTHTDISTKD-----ENHNDKINNGVINI 666

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 2, 2003, 16:14:11; Search time 52 Seconds

(without alignments)
2382.024 Million cell updates/sec

Title: US-10-006-780-2

Perfect score: 6842

Sequence: 1 MNSKIKVVRKRPISLEKK.....KKLVQDNKSNMHNHKK 1288

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	789.5	11.5	4550	2 T18440	hypothetical prote
2	778	11.4	1844	2 D71612	hypothetical prote
3	765	11.2	1619	2 T18499	hypothetical prote
4	755	11.0	668	2 T02791	mitotic centromere
5	752	11.0	3394	2 T18501	hypothetical prote
6	748.5	10.9	1817	2 D71606	hypothetical prote
7	732	10.7	1254	2 T18277	kinasin heavy chai
8	728	10.6	716	1 A44259	kinasin-related pr
9	722	10.6	671	2 T10755	kinasin-related pr
10	698	10.2	1436	2 D71618	hypothetical prote
11	696	10.2	807	2 B71605	hypothetical prote
12	676	9.9	2150	2 S71629	senesory transducti
13	656.5	9.6	747	2 T23607	hypothetical prote
14	646	9.4	1256	2 S14556	asparagine-rich pr
15	611	8.9	1711	2 T18429	hypothetical prote
16	603.5	8.8	2523	2 T18477	hypothetical prote
17	597	8.7	4981	2 T18489	hypothetical prote
18	591	8.6	537	2 A23770	asparagine-rich pr
19	587	8.6	3724	2 T18427	hypothetical prote
20	584	8.5	2485	1 H71621	serine/threonine-s
21	581	8.5	1650	2 T18444	hypothetical prote
22	579	8.5	2708	2 T09079	probable chloroqui
23	579	8.5	3844	2 T18402	asparagine/asparta
24	577.5	8.4	2380	2 E71604	hypothetical prote
25	569.5	8.3	2500	2 G71609	hypothetical prote
26	567	8.3	2010	2 B71616	phosphatase (acid
27	563.5	8.2	2819	2 T09080	probable chloroqui
28	561.5	8.2	3973	2 B71612	hypothetical prote
29		8.2	1111	1 A42640	kinasin-related pr

30	560.5	8.2	2339	2 A45597	DNA-directed RNA p
31	552	8.1	2539	2 B71619	hypothetical prote
32	549	8.0	2510	2 T28160	hypothetical prote
33	544	8.0	2269	2 T18472	hypothetical prote
34	543	7.9	1984	2 A44396	P-type cation tran
35	542.5	7.9	1553	2 T18502	hypothetical prote
36	541.5	7.9	1073	2 S14032	kinasin-related pr
37	538.5	7.9	1085	2 T38378	kinasin-like prote
38	537.5	7.9	1714	2 E71609	Ser/Thr protein ki
39	534.5	7.8	1516	2 E71619	RAD endonuclease
40	534	7.8	1058	2 T47525	kinasin-related pr
41	533.5	7.8	2391	2 T18410	cardamoyl-phosphat
42	533	7.8	1390	2 T14004	trifA protein - sal
43	532	7.8	813	2 T46242	kinasin-like prote
44	529.5	7.7	669	2 S14535	asparagine-rich pr
45	528	7.7	1166	2 H71609	hypothetical prote

ALIGNMENTS

RESULT 1

T18440 hypothetical protein C0425w - malaria parasite (Plasmodium falciparum)

C/Species: Plasmodium falciparum

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 09-Jun-2000

C/Accession: T18440

R/Lawson, D.; Bowman, S.; Barrell, B.

Submitted to the EMBL Data Library, August 1997

A/Reference number: Z18935

A/Accession: T18440

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-4550 <LAW>

A/Cross-references: EMBL:Z98547; NID:e1325376; PID:e1325396; PIDN:CA11121.1

C/Genetics:

A/Map position: 3

A/Note: C0425w

Query Match	Score	DB 2:	Length
Best Local Similarity	21.0%	Pred. No. 2.7e-25;	
Matches 380;	Conservative 247;	Mismatches 498;	Indels 683; Gaps 70;
QY	11	KRPISLEKKKSDIIVKNNCTLYID-EPYKVDMTYIRHBFIVDKVDDPTVD-NF	68
DB	795	KQDDNKDDKDDDD---KQDDKYDDGDDKTYDDDDNKYDDDD---DDIYDDDDDFNF	845
QY	69	TYENTIKPL-----IIDYENGCV-----CSCFAYGQ--TSGKTYTW-----L	106
DB	846	DHDSKSKLSNVSDILKDKKGGSLKEDTLKSGKMINVQKSKSKMYCMTNNCVV	905
QY	107	GSQPYGQSD-----TPGIFYAAGDIFFLAIYDKONTKGFISFY	147
DB	906	KEKNYFVNDYMHINLSDHEMNIMMSPG--YNNLFDVDFHFKRD-----FI	954
QY	148	EYCGKLYDLQRRKVAALENGKEVVYKDLKIRVITKEEL-----ILKMI	195
DB	955	NANILKLT---KEKIIQVLQNNYKQVCYFCGKTGKTKCMFBCSITYFHYCYNNKM	1010
QY	196	DGVLLRKIGVNSQNDSSRSIALINLIDLKINKNTSLGKIAFTIADGSRGADVSONKQ	255
DB	1011	QVRRKKNDLQASQKSR-----KNGRAASV-----KRGAE-----GK	1045
QY	256	TQTDGANINRSLALKECIRAMDSKNHIFPDSLTLYKVLIDPVGKSKSIMIANISPTI	315
DB	1046	DKKDS-----RSIDRRKKDKVTRRSYESSANDSDKKNNG-----	1082
QY	316	SCCEQTLNLTARYSRVKNFKNKSTCINEDDTNTRIRIILDSKSGEMNAASITENVIKSN	375
DB	1083	-----DDKKKGDIKKNNNNNNNNDSDSNNNDSDSNNNDSDSNNNDSDSNNNDSDS	1135
QY	376	HLLSNNNNKIN-----GKINDKIERNNIIKNSFDKP	409


```

QY 556 -----LKDNIKLNRSSCDINIMKKNNLHARHVSGLTMTFSDPO----- 598
Db 426 SFHDLPLNMQVLAKNVVMQMEHNF-IVMDEMKNVNVK-----RINSISHIYMANLKN 479
QY 599 -----KKKNDTPFKSNINKEDETPKDI---LYESRNVSNNGNVILG-LNKOTHH 646
Db 480 YNEHTSMVKGKGTNRKKSNNLKIINNPGSVNARAISENNQSTPAHGNI PMGSVDKYL KHD 539
QY 647 ISFKD-----ENNDKIN-----NGVINI INNSVNSIN-----SMMNSINN-SNM 688
Db 540 RMDNDLKNINNNNNNSMNNNSMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 599
QY 689 NSNSIYKSNYSNQSII-----SDVQIRYVNEMDTSK-----NND 723
Db 600 NNTSILN-----NNKKITRGRAKKSTINININIKNSSTNKSMTNNSVNNMNSVNM 656
QY 724 NTFEDALSCDN-NMPTITNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 780
Db 657 NSYKNNNSVNNNSVNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 716
QY 781 HNLF-----QPDNNKNTSNIQNTN-----TNKNQOGNVVYSNNFCHYNL 820
Db 717 NMFNVNPNQNLINGIMKDIINNNTITVSNKKLMNNYINDNNIMMESINETYNF-DGTL 775
QY 821 NDKVYLIDLNNKEQKDNHIGCDN-----NIIQNRNDFEKKKKTNFY 862
Db 776 NNKK-VSNNNNNDIHDGVLNTLRSSSYIKPHRLTLPIMYNNNTYNNYSKAYNY 834
QY 863 NNNNVIV-----NNNKNNSSPMKYGLCGSHTSIDMKNNEMKNNEMKNNH 917
Db 835 ENQNNMAQODCDNKGSLNEGPRYN-----DNNNNNNNNNNNNNNNNNNNNNN 883
QY 918 SNNNSSSSSSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 966
Db 884 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 943
QY 967 NIY-----QNDIIYITINSIDYMSN-----TLL 990
Db 944 NMVDKAAVKNMMLIGHSNYDLMKNGMNGIKLMOEVQKMGNEVGKRAARRAKTI 1003
QY 991 HFEKTYPTLSTNEDIYNKMEGKHIRLDDQKYDNDN-----NNVD 1034
Db 1004 NLKSIRFNTVSLG-----LKGNH-NMENNNMMNDMMNNNNNNNNNNNNNNNNNN 1055
QY 1035 N-NKKNNVD--NNVDN--NNVDN--NNVDN--NDKNNVDN-----NNVDNDDVDDEHNK 1083
Db 1056 NVNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1115
QY 1084 NNEYLAFQKN-----VDTIINCLNSLIDISSMYDDTKEILNNILLSKYKAKONV 1136
Db 1116 NMRMRMHIDNNMTYNTLNMKNKINNAGNNLN-----GEFSKILYN----- 1156
QY 1137 KYINEDIKNNLSIEDKTAQSIYERKVLJTKLLLFKKNNVDTOINNETSD-----LRKD 1191
Db 1157 -----RSKSTENIKEMQPHNEANNI-----SGNNTSDSNITLTKN 1193
QY 1192 LVWCHICNNNP--DDQFH-----FYAGRLKEDIINLMLRQICE-----SENRL 1238
Db 1194 LVEMIRNNKPSIDKIHESVNSYSPLMNIGSSYLKKKKTAEIKTGENNTENNKGIV 1253
QY 1239 QPLVYKONKSANSVLLNV-----SSNNGDIILNKKLVQDNKKN-----SMD 1281
Db 1254 VNSQVEKSGSEKSIIEIKEGQIKVLYNNNNNNNNNEKDQADQYODONKHKHQDOR 1313
QY 1282 HNNIHK 1288
Db 1314 QNKSHQ 1320

```

RESULT 3
T18499
hypoetical protein C0770c - malaria parasite (Plasmodium falciparum)

```

C:/Species: Plasmodium falciparum
C:/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:/Accession: T18499
R/Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A/Reference number: Z18935
A/Accession: T18499
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1619 <Law>
A/Cross-references: EMBL:Z88551, NID:e1331903, PID:e1331920, PIDN:CA11138.1
C/Genetics:
A/Map position: 3
A/Intons: 87/3
A/Note: C0770c

Query Match 11.2%; Score 765; DB 2; Length 1619;
Best Local Similarity 22.0%; Pred. No. 9, 6e-25;
Matches 369; Conservative 246; Mismatches 568; Indels 496; Gaps 66;

QY 5 IKVVRKRLPELEKK-KDSIIITVKNCTLYIDPRYKVDMTYIERHEFVDPKVPD 63
Db 16 IKVVRKRLPELEKK-KDSIIITVKNCTLYIDPRYKVDMTYIERHEFVDPKVPD 69
QY 64 TVDNPVYENTIKPLIIDYENGCVSCFPAYQSGKTYTMLS-----QPYG----- 112
Db 70 DVDQKTLFNNYIQ-IVDEVLOGFNCTLFYCGQTGTGKTYTBGKILHKLQYDNKKYD 128
QY 113 -----QSDTPGIFQYAGDITFL-----NIYD-KDN-- 138
Db 129 LNESINSDISYCYELCENEDTGLIFRVTRKIPDIINLKRKEEKIRHFKKKNIFDEKDNH 188
QY 139 -----TKG----- 141
Db 189 INVINKETNGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 248
QY 142 -----IFISFEIYCGKLYDL-----QKKNVAAAL 167
Db 249 VVQKSVDRKNNNDVQTFDDKMYDPFNKISYLEIYNEELCDLSSSTENKKLRYEDS 308
QY 168 ENGKKEVVVKDIKILRVLTKEELILKMDGVILRKIVNSQNDSSRSHAIINIDL--KD 225
Db 309 NKSQGLNDKLEBSKINSFEIYIYICAIKKRRTAEAVYKSSRSISFTITLID 368
QY 226 IN-----KNTSLKIAFIDLASERG-ADTVSQNKQOTDQGANINSLAKCICIAMDS 280
Db 369 INNVGESITIKGLNVLDAISENALKSSYGLKTRQDSQINOSLTLGKIVNSLLEN 428
QY 281 KXHIPRDELTFRVLDIFVGRKSIMTANISPTISCEQTLNTRYSGRVKNFRKSTYC 340
Db 429 SSYIPRDSKLRLLQDSLGKTKTFIYATISPSLCLDETISTLDYVRANKINRP-- 486
QY 341 INEEDTNTERTISILDSKSENNASSIENVVIXSHNLLSNNNNNKINNGKINDKIERNI 400
Db 487 -EINKTKTKQKIKD-----LNNEIKELNAL--NLSEKSGVYLDTEEVYNI 531
QY 401 -----LKNS-----FDRPREGTSTFGYSSLNDI-----DKIKKKK-- 434
Db 532 QNSLKKKEILLEKEKILFEKSK-IKULLNKQVTDLDLQNKIFFLKDTQKYXNQIAF 590
QY 435 -----KGLINYKSTLYNDVITNKKNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 484
Db 591 HQILIKITIEEK--YITQFLANKYSLEBNFNQCKSSFOQITMNSNIMHDLNLKNI 648
QY 485 NMINEMINNINNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 536
Db 649 HOENHLNDICTIILFIIDTQVYLVQKKSYNDYINHFEBELNSINIKYNDTLNF-IID 707
QY 537 DNNCHLNNDKSIFLHKXLR-----DNKIKL----- 563
Db 708 HMDI-LQKYDHDYFSLSENIEKKEISQCKINIPHISDEAKQKPAIIGVYNTNGLKEKOS 766
QY 564 -----NRSKCDINNNKKNNLH-LARHVSGLTMTFSD----- 596

```


Db 767 G E K I T C K O S G E K I T C K O S G E K I T C K O G E I T C K O S G E K I T C K O G D E I C E K O N 826

Qy 597 -----DQKAKDNTFFKSNINKEODN--TPKDIYESNV--SNNGVALLGLAKNTHTDIS 648

Db 827 G E G I I C K O N G E K T C Q K O Y K G E E N N E M P F H I E S F N Y P E N K E F I K D I L L A N O G E D I I Q 886

Qy 649 T K D E N H A D N K I N N -----G V I N I -----I N S N 671

Db 887 N L D N H E S E K K L Y A N K S L I O Y I K S I C N I F H I F L N E S F M L E R K C H E K Y I H W D K K I M N D N 946

Qy 672 V-----NSINNSNNMSINN--SNMSNSIYKSNYNYSNOSISDVOIRYVEM 715

Db 947 I E Y I Y K K Y D V P E K E D N I I I N N Y K K I M E D I O E F K E N I L K O I E L L I N G A V E D I Y E K A N K F 1006

Qy 716 -----D T S N K N D N I F P D A I S C D N N M Y P N I T N N N N N N N N N N N N N N N I D E V E N Y N R D 767

Db 1007 D L A N K E L Y D K S K K Y K V F M N S Y K K I D T I Q I H F S E C I N D E K E Y N I N N D I -----K R K 1060

Qy 768 G T N N S M K L Y A N S H -----N L F O P D N N K --N T S N I O N --I N T Y K A N Q O G V A N Y S M F C H 817

Db 1061 I T N H T V L Y N I L S Y L K R S Y L E E K E S K C L A N D T I N I Y N N L Y T M I N I A D I K I N -----1113

Qy 818 Y N L A D K Y V L I D L N --N K E O K D N I G C D N N I I O N N D E P E K K K T N -----F Y A N N N I Y I V N 871

Db 1114 -----N F T S O T N I N K I E K R K L Y N V I S E K I I O N K S O E O K G G R R I H M L K S N E N I G F N 1167

Qy 872 N N M G N N N S P R A K Y G L C G S H T S I D M K N N K M E K N E M K N E M K D N H I K S N N N S S S S S S N 930

Db 1168 N I F E -----Q V L N K N M Y I A L I D I L A K K O Y L O O I P S T S S D L M E M N 1209

Qy 931 --N N I Y N N I N D -----D T F O N -----D Y C H A N D T P I R K K N T N I N S N I Y O N D I 974

Db 1210 D D E F Y L N I K D V O R I R E T F V N V I D M T P M K I T S I A N R D S T Y -----D T N Y D N N --K N D D 1260

Qy 975 I Y T I N S L A N D Y M S N T L L P F K E K Y T P T L S T E D I Y N K M E G K I R L D O D K Y D D N -----1028

Db 1261 -----N N K N N H N N K O N H D N N K N H N -----D N N K G G N N K N D H N N K G N N N Y I D D K A D O G H 1313

Qy 1029 D N N N V D N N N K N N V D N N Y D N N V D N N V D N N --D K N N V D N N N V D N D D -----D V Y D F H N I K N F N N 1084

Db 1314 D H N N Y L V I N D D K M D O G H D H N N Y L Y N D D N K M D O G H D H N N Y L Y N D N D K M D O G D H N N Y L Y N D 1373

Qy 1085 N-----E Y L S F O K A N V D T I-----I N N C L -----N S L D I S S M Y D -----1113

Db 1374 D N K I D O G H H N Y Y L Y N D D I I R O N L Y L E V K N T L M K E I B O I T T T H D I D S K F E O G I D E N F O F 1433

Qy 1114 ---D T K E I L N N I L S K Y A E K D N V I K K Y I N D I C N M S L E E I D T A O S I Y E R K V L L R K L 1170

Db 1434 L I O D K I V A H E N N Y S R Y N S L A D H T T N S S I P H N I N K G --D K R K L V A K T N H N N L P L N E Y 1489

Qy 1171 L L F K K A N D T O I N N E T S D L R K D L A M C H I C N N P D O Q F H Y A S R L E K D I I N L --I M L R O I W C 1229

Db 1490 M L Y N K M T ---I D N I P I O P I D E T Y O E L A N Y I N N O H N K -----R D V A T E M N S V S L A K E Y E 1539

Qy 1230 E S E N I R L L Y O F L V E Y O N K S A N S V L A N V S S N G D I I L N K K L V O D N I K S M D H N N I H K K 1288

Db 1540 R K R I N V O N V A I T N N Y S N N N N N N N N N N N N N N N N N -----N N K F O S T K I K -A I E N R A P F E K K 1593

RESULT 4

T02791

mitotic centromere-associated kinesin MCAK [imported] - Leishmania major (strain Friedlin

C:Species: Leishmania major

C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000

C:Accession: C61455; T02791

R:Wyley, P.; Audlemann, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.;

A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c

A:Reference number: A61455; MUID:99176987; PMID:10077609

A:Accession: C61455

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-668 <PLY>
 A:Cross-references: GB/E001274; NID:e13264850; PIDN:AAC24615.1; PID:g2978452; GSPDB:GNO
 A:Best Local Similarity 42.9%; Pred. No. 9.8e-25;
 A:Experimental source: strain MHOM/IL/81/Friedlin
 C:Genetics:
 A:Gene: MCAC
 A:Map position: 1
 A>Note: MCAC
 C:Superfamily: kinesin-related protein KIF2; kinesin motor domain homology
 F135-459/Domain: kinesin motor domain homology <KNOT>

Query Match 11.0%; Score 752; DB 2; Length 668;
 Best Local Similarity 42.9%; Pred. No. 9.8e-25;
 Matches 172; Conservative 67; Mismatches 134; Indels 28; Gaps 8;

```

Oy      3 SKIVVVRKRPLSLSELEKKKDDIIIVKNNCTLIYDEGRVYDMTKYIERHEFIYDKYED 62
Db      133 SRIIVAIKRPRLSAGEBOTNGFTDIMDADNDSSEIVLKEPKVADLRKYIHVHRFFFEYDFD 192

Oy      63 DTVDNFTVYENTIRPLIIDLTVENGCVCSFPAVGQTSKTYTMLSGOPVGSQDPPIQY 122
Db      193 EACQNVQVYNAARALIDYVDPGCG-ATCFAYGQTSGSKTHMYLKGK-EBGLVAL 246

Oy      123 AAGDIFPFLNYDKONTKGIPTISYEYICGLTYDLLQKKRVVAALENGKKEVYVNDLKITL 182
Db      247 AAKMPFRLT----SDTR-IYVSFEYIYSGTLPLDLNRRRLRALIEDDKGRVINGRLTEH 301

Oy      183 RVLKKEELIKMIDGVLLRKIGVNSQNDSSRSHALNITDKINKNTSLGKIAFIDLAG 242
Db      302 CSTVEDMTITDQSGSGSCSTGANDTSRSHALLETIKLAKRTSQSGKFTIIDLAG 361

Oy      243 SERGADTVDSQNKQOTQDGCANINRSLALKECIRAMDSQKNIIPERSELTQVLRDIYVGK 302
Db      362 SERGADTVDCARQTRLBGAIRKSLALKECIRFLDQKRKHVPEFGSKLTVLRDSFIGN 421

Oy      303 SKSLMINIAISTICCEQTLNTLYSSRVKRFKXST-----CINEED-----DNTIR- 351
Db      422 CRTVMIGAVSPSNNAHEFTLNTLRYADRVKELKRNATERRRTVCPWDDQDEAFDTTESRP 481

Oy      352 -----ISILDSKSEMNASSIENVVYIKSNHLSNNNNKKN 387
Db      482 PSRRITTRLSTAAPLFGSGSTAAPALRST-LISRSVNTLS 521
  
```

RESULT 5
 T18501
 hypochemical protein C0760C - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C:Accession: T18501
 R:Lawson, D.; Bowman, S.; Barrell, B.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: Z18935
 A:Accession: T18501
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-3394 <LAW>
 A:Cross-references: EMBL:Z98551; NID:e1331903; PID:e1331922; PIDN:CAB11140.1
 C:Genetics:
 A:Map position: 3
 A:Note: C0760C

Query Match 11.0%; Score 752; DB 2; Length 3394;
 Best Local Similarity 21.9%; Pred. No. 7.4e-24;
 Matches 308; Conservative 249; Mismatches 433; Indels 416; Gaps 58;

```

Oy      15 SELEKK-----KKSDIITVKNNTLYIDE---PRYVDMTKYIE 51
Db      2187 SELEKKNDMLENYKNELEKNEBEIYKLANDNDMLNNCKKLESIMMEKYKIIINNNTIQ 2246

Oy      52 RHEFIVDKV-----PDQTVDNFTVYENTIKPLIIDLTVENGCVCSFPAVGQTSKTYT 104
Db      2247 EKDELTEENKKNKKNKKDLNLNNISVVDKSI-----VSCF----- 2281
  
```



```

QY 105 MLAGSOPYGSDTFCGIFGYAAGDIFFTLINTYDKONTGIFISFEYIYCGKULYDLQKRW 164
D 2282 -----EDSNIMSPSCDIDLNFVNNLSKSNK-----VCTMMDICENMDDSSISIN 2328
QY 165 AALENGKEVAVVVDKLTFLVLTKEELFKMIDGLKIGVNSOND-----211
D 2329 NNINNVNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2387
QY 212 -SSRSHAILNIDLKDNKTSILGKIAFIADLASERGADTV-----SONKOTOTGANINR 265
D 2388 LGSGLYLTIRNKLKEI-----QMLKNQILSLBESIKSLNEPFLNKLAKNEKNEKELLIKIN 2443
QY 266 SLALKECIRAMDSD--DKNHIIPRDELTKV--LRDIFVGSKSXIIMANISPTISCC-- 318
D 2444 EILKTKNNLQDNESCIONLNNYLLKKNELNKNIVKNIF--KYGYITLILQGSVFCIF 2501
QY 319 -----EQLTLNLRYSRVKNF-----KAKSTCINER--D-----TN 348
D 2502 KHFENKTIIDQSLITNKLTLKSPDFYWDVSVIQEIENKKNITIIIDQFLTDEYFKAIQTF 2561
QY 349 TERISILDSKG-----SENNASIEVNVIKSNHLSNNNNNNKINRGKINDKIERNNILK 402
D 2562 TKTCNVILQRCYLSILKDTNNDFIQN--KQSNQOQNGNHNIMNCNITPDBEINVTAD 2618
QY 403 NKSFPDREGFTSTFGKYSLNDIDIKIKKKKKGLINYSKTYNDTIN--KKHNNNNN 460
D 2619 QQIFD-----GTENVQOQLQNEBEDYVNEEM--YTDKMDLNNMNGDDDDDDDD 2667
QY 461 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 520
D 2668 DDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2726
QY 521 OPIYAFDTDSFLSDMNCNLNNNDKSIPLAKKKLDNFKLN--RSSCCNINMKKKNNL 579
D 2727 -QNIYNDVNDNNV-----INIKILIKKODIINDLCKOEHNNILRLINSEIENA 2779
QY 580 HLAHSHVSGKLTWESYDPOKXKNTPEKSNINRQ--EDNTEKD--ILY--ESRNVSMNGN 634
D 2780 H-----NLENVYQVNDANLINDINIKKEFLTFVDEKDVSN--2816
QY 635 VILGLKNTHTDSTKDEHNDKINNGVNIINNSVNSINNSNNNSINNSMNSISY 694
D 2817 -----ESNSKCD-----DDKKEMDIIOAKKENPVASTHYDNDNDIKDIN--2858
QY 695 KSNVNSQSIDVQIRVNEMDTSNKKNDNIFPDALSCDNNVYPTITNNNNNNNNNNNN 754
D 2859 --NDNNNDNIND-----DNNNDNI-----NNNNNDNINDN 2888
QY 755 NNNIDVENYNNRDGTNNSMKLYAVNSHLEFQPDNNKTSNIQNI--NTKKNODGNVYSM 813
D 2889 INNDNINNDNNNDNNND-----NSNNGFVCELSSTINDFNILNVNKNFQ--INKSN 2940
QY 814 NFCHYLNLDKQY-----LID-----LNKEQDKNHCDDNIIQNRNDEKKKKKTYNFYNN 864
D 2941 NF--STNLENYAYAVYKIVEAGSALENKKOD-----KKRKYFSD 2979
QY 865 NNIVYVNNNNNNNSPMMKYGLQGSHTSIDNKKNNKNNKNNKNNKNNKNNKNNNS 924
D 2980 SETTKENGYLGESNSIKIRKYSVVG-----DNDNNNDNDNDNDNDNDNDNDNDNDND 3035
QY 925 SSSSS-----NNNINNNINDDTFQNDYCHNDNTFTIRKKNNTINSNIYOND 973
D 3036 DNDSDSYDEKEQENSDNVTELNHVTEDISVHN-----IQENKKKYFENVHENTF 3090
QY 974 IITINSINDYMGNTLHFEKXTYPTLSTNED-----YINKMEGKILRLDDQ 1022
D 3091 NLAHVEDVHEDLQ-----YEDYVNNNNNNNDNDNEESNNSCYIISDDEGKKNVSKQ 3145
QY 1023 DKYDNDNNNNVNNNNKNNVNDNNVNDNNVNDNNVNDNNVNDNNVNDNNVNDNNVNDNN 1082
D 3146 NDDDEDDDDGDDDDNEDDDDDGDDDD--DNEDDNDGDD--DNEDDNYDEDEDDNY 3202
QY 1083 NNNEVLSYFOKNVDIIINNCLNSLIDISSMYDTEKILNNILSLKYAEKDNVILKYINED 1142

```

```

D 3203 ENEDDDNVANEDDDNYEN-----DDDDNYENDD-----DNYENDDN-----YEND 3245
QY 1143 IKKMSLEEDKTAQASIEYERKVLTLTKLLLFKKNVDTQIINNETSIDKRLVWCHICNNP 1202
D 3246 DDN-----YE-----NGDDNDNDNDND-----DNDND 3266
QY 1203 DDQHFYAVSRLEKDIINLIMLRQIWCESENLRLLYQFLVVEYONKSANSVLLNYSNNG 1262
D 3267 EEKYS-----CHND-----KQEHNTNDLNLIDHDN--3291
QY 1263 DITLKKLVQDNIRN-----SMDHN 1283
D 3292 -----NKNNTDELSTYVNSVSHN 3311

```

RESULT 6

D71606
 hypothetical protein PFB0800C - malaria parasite (Plasmodium falciparum)
 C/Species: Plasmodium falciparum
 C/Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
 C/Accession: D71606
 R/Gardner, M.J.; Tectelin, H.; Carnucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.,
 .; Perera, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
 Science 282, 1126-1132, 1998
 A/Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
 A/Reference number: A71600; MUID:99021743; PMID:9804551
 A/Accession: D71606
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-1817 <GAR>
 A/Cross-references: GB:AE001418; GB:AE001362; NID:g3845275; PIDN:AACT1949.1; PID:g384521
 A/Experimental source: clone 3D7
 A/Genetics:
 A/Gene: PFB0800C

Query Match 10.9%; Score 748.5; DB 2; Length 1817;
 Best Local Similarity 23.6%; Pred. No. 5.3e-24;
 Matches 356; Conservative 197; Mismatches 410; Indels 543; Gaps 69;

```

QY 79 IIDLYENGVCVSCFAYGQTGSGKTYTMLGSPYGSDFPGIF-----QYAGDIFFTLNI 133
D 391 IELFTEGCVLFLDNEKIKSNIN-----TPIFLKLLKNTYTSNDFV--I 435
QY 134 YKDNTKGIFISFEYIYCGK-----LYDLQKRWVAALNGKEVAVVVDK--- 180
D 436 HINDRLNRYIKFLSYHGSNNIYVHILENDLYTHKKIF-----NNYDMLIKHYKSTA 490
QY 181 --ILRVLTKEELFKMIDGLVLRK--IGVNSQDNDESSRSHAILNIDLK-----DI 226
D 491 DHILVLTNKG--NLWYLVNTILLNNYSLYNKKENNEKQSLNKLKVPFSLHYNKQHTT 549
QY 227 NKQTSILGKIAFIADLASERGADTVSONKOTOTGANINBSLALKECIRAMDSDKNHIF 286
D 550 NKNVNV-----DPNQOHKHXDKD-----SYDNTYEQMKNNKKKIYP 584
QY 287 RSELTFLVLRDIFVGSKSXIIMANISPTISCCBOTLTLRYSRVKNFKNKSTCINEED 346
D 585 NRYITTHILQ-----NYEONLYSPQK-----DD 609
QY 347 TWTERTSILDSKSENNASIEVNVIKSNHLSNNNNNNKINRGKINDKIERNNILK----- 402
D 610 TWINNIFDL-HKREGDIYEKEN-----ESSDIIPRDSYKKKIKKEK-----KKKNITKYEDH 660
QY 403 --NKSFPDREGFTSTFGKYSLNDIDIKIKKKKKGLINYSKTYN-----DNTINKHNN 457
D 661 PLNKEKKKKKKKFFYINYEKGDKNDNDL-----YANNIYSKQLENIQNKQYVNN 709
QY 458 NNNNNNDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 517
D 710 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 564
QY 518 HLPPIYAFDTDSFSL-----DNMCHLNNDKSIK--LHKKLLRDNIK-----LKN 564

```



```

QY 793 -----SNIQNTNKNODGN-----VNSMNFCH-----817
Db 654 NDLAKENGGQQQQQHNDDEDEIKSTSNATTTTITIDMASHSGIDPIEFTLIK 713
QY 818 -----YNLDKNYLIDLNKQ-----KDK 837
Db 714 SDKTTTIERETIQPSSLSNSTSLDIEFVESSTLPAPPVTTTITITVTTKLTKT 773
QY 838 NHCQDNNTIQNRNDFEKKKTNNYNNNVIYVNNMGNNSPBMKGLGSGHSIDNMK 897
Db 774 NIPS-NTNDINSIDDFGFSK-----IEEGSSSNKKP-----NDTALLSFG 813
QY 898 NNEKKNEMKDNKDHKIKSNNSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 953
Db 814 DDDDEENE--DNEDEDYVDSDE--TISGKNLLN-----TFNDHHRGDFGAPT 861
QY 954 FTTRKN-----NTNINSNY-----969
Db 862 KSIFNKNNTITKEFEFPQOQQQQQQQQQQQQQQQLIQTSTSTNFTIISKSNKE 921
QY 970 -----QNDIITYTNSLNDVNSNTL-----HFEKXT-----997
Db 922 PPSSTSTTTSIKKKKNFKRSWITFTLTITVSSSLCLCYPEYERLVRGRGYLTKL 981
QY 998 --YPTLTSTNEDI--YN-----KEMEGKHIR--LDQDQKYDDNNNNNNNNNN 1044
Db 982 GYSDYFTNEKISLAQNOISLAKELYGKSKOYDEMSFNTAYNHLIEMHLETAISK 1041
QY 1045 VDNNDVNDVNDNDK--NNVD-----NNVDND--DDVDENIKNNNNNEYSYFQ 1092
Db 1042 IFPSAIDLRFGSDVINSICKRAIHKLKTNNYVGNLDQOQQQOHNY--ITKIDQSL 1099
QY 1093 KAVDTIINNCLNSDISMTDPTKEILNNILSKYAKKNVKKYIN--EDIKMSLEB 1150
Db 1100 KEQQLIENF--KLDLK--NKTSELEK--LEKEIKQKNKKEKEKIELLESRYLNE 1151
QY 1151 IDTKAQSIYE-----KRVLLTKL-----LTFKQVND--TOINNETSDLRK 1190
Db 1152 -EKGEKVLVDQIISLNDKNTLSTQILNEGDKSLGVLYIKLNSKTEIQNEVELEK 1210

RESULT 8
A44259
Kinesin-related protein Kif2 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 19-Jan-2001
C/Accession: A44259; S27871
R/Alzawa, H.; Sekine, Y.; Takemura, R.; Zhang, Z.; Nangaku, M.; Hirokawa, N.
J. Cell Biol. 119, 1287-1296, 1992
A/Title: Kinesin family in murine central nervous system.
A/Reference number: A44259; MUID:93077686; PMID:1447303
A/Accession: A44259
A/Molecule type: mRNA
A/Residues: 1-716 <AIZ>
A/Cross-references: EMBL:DJ2644; NID:g220467; PIDN:BA02165.1; PID:g220468
A/Experimental source: brain
A/Note: sequence extracted from NCBI backbone (NCBI:P:118908)
C/Superfamily: kinesin-related protein Kif2; kinesin motor domain homology
C/Keywords: ATP; microtubule binding; nucleotide binding; P-loop
F:196-531/Domain: kinesin motor domain homology <KMOT>
F:285-292/Region: nucleotide-binding motif A (P-loop)
F:291/Binding site: ATP (Lys) #status predicted

```

```

Query Match 10.6%; Score 728; DB 1; Length 716;
Best Local Similarity 47.6%; Pred. No. 1.4e-23;
Matches 161; Conservative 61; Mismatches 102; Indels 14; Gaps 8;

QY 4 KIKVVRKRPLESEKKKSDITVKNCTLYIDERYVDMTKYERHEFLVDKFPD 63
Db 195 RICVCVRKRPLESEKKKSDITVKNCTLYIDERYVDMTKYERHEFLVDKFPD 254
QY 64 TVDNFTVYENTIKPLIIDLYENGVCSCFAYGOTGSGTYTMLGS-QPYGSDTPGIFQY 122

```

```

Db 255 SAENWEVYFTAPLVEITFERG-MATCFAYGOTGSGKTHMGDGSFKQDSCKGYAL 313
QY 123 AAGDITFTNLYDKMTK----GIFISFEIYGGKYLDLQKKRMVAALENGKEVAVKD 178
Db 314 AARDVPLMK--KPYKKLEIQVATPFETISGKVPDLNRTKRLVLEDGQQVQVVG 370
QY 179 LKILRLYTEKEELIKMID-GVLLRKIGVNSQNDSSRSHALINDKIDINKTSLGKIAF 237
Db 371 LQREKVCYED-VLKIIDIGNSCRTSGQTSANAHSSRHAVPQIILR--RKGLHKEFSL 427
QY 238 IDLAGSEKADTVSONKQOTGDGANINRSLALKECIRAMDSQKNI PRDSELTQVLRD 297
Db 428 IDLAGNERGADTSSADRRQRLREGAEINKSLRLKECIRALGRNKPPTPFRAKLTQVLRD 487
QY 298 IFVCK-SKSIIMANISPTISCCQTLNTRYSRVKNF 334
Db 488 SFIGENSRCTCMATISPGMAASCENTLNTLRANRVEF 525

RESULT 9
T10755
Kinesin-related protein 2 - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 29-Sep-1999
C/Accession: T10755
R/Sperdy, A.O.
submitted to the EMBL Data Library, January 1998
A/Reference number: Z17121
A/Accession: T10755
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-671 <SPB>
A/Cross-references: EMBL:U44979; NID:g2772515; PID:g2772516
A/Experimental source: strain Sprague-Dawley, testes
A/Genetic: KRP2
C/Superfamily: kinesin-related protein Kif2; kinesin motor domain homology
F:205-542/Domain: kinesin motor domain homology <KMOT>

```

```

Query Match 10.6%; Score 722; DB 2; Length 671;
Best Local Similarity 41.2%; Pred. No. 2.3e-23;
Matches 171; Conservative 78; Mismatches 138; Indels 26; Gaps 11;

QY 4 KIKVVRKRPLESEKKKSDITVKNCTLYIDPRKVKVDMTKYERHEFLVDKFPD 63
Db 204 RICVCVRKRPLESEKKKSDITVKNCTLYIDPRKVKVDMTKYERHEFLVDKFPD 263
QY 64 TVDNFTVYENTIKPLIIDLYENGVCSCFAYGOTGSGTYTMLGS-QPYGSDTPGIFQY 122
Db 264 TASNEVYVFTARPLVOTIFBEG-KATCFAYGOTGSGKTHMGDGSQNSKGIYAM 322
QY 123 AAGDITFTN--LYDKNTKGIIFISFEIYCGKYLDLQKKRMVAALENGKEVAVKDK 180
Db 323 ASRDVPLLNQPRYRSLNE-VYVTFEIFYNGVFELLNKAKLRLVEDSKQOVVVGQ 381
QY 181 ILRLYTEKEELIKMID-GVLLRKIGVNSQNDSSRSHALINDKIDINKTSLGKIAFD 239
Db 382 EYLVTGAD-VIKIMNGSACRISGQTFANSSRSHAFQIILR--AAGRLLHKEFSLVD 438
QY 240 LAGSEKADTVSONKQOTGDGANINRSLALKECIRAMDSQKNI PRDSELTQVLRDIF 299
Db 439 LAGNERGADTSSADRRQRLREGAEINKSLRLKECIRALGRNKPPTPFRAKLTQVLRD 498
QY 300 VCK-SKSIIMANISPTISCCQTLNTRYSRVKNKSTCINEEDTNTERTISILDSK 358
Db 499 IGENSRCTCMATISPGISCEYTLNTRLYADRVKELSPHSGSGEG-----AVQME 549
QY 359 GSEMANSENVIVSNHLSNNNNKIN--RGKINDKIERNNILKNKSPDKRE 411
Db 550 TEEMDSS-----HGASLTGNEEELSQMSFNEAWTQIRELEERAMEURE 597

```


Best Local Similarity 26.6%; Pred. No. 3.5e-22;
Matches 253; Conservative 151; Mismatches 300; Indels 246; Gaps 43;
QY SSRVQPKKSTCINNEEDPTNTERISLDSKSGPMASSTIENVYKSNHLNNNNKIN 387
D 10 SSKKEEEN--DIINKCDSSN-----KINGKE-NFAVEKGINESGMSDNIKN-N 58
QY 388 RGIINDKIERNNILKNKSPKREGFTSGKYSLSLNDIDKIKKKKGILNYKSTLYN- 446
D 59 QEKKKKKKKKKHKKYKYNNTHTNHTNDKNGG-QDIKPEYEREDNINIKNDPNNI 117
QY 447 -DNTINKKNNNNNNNDNNNDNNNNNDSSSVNN-----MINEMINNI----- 495
D 118 LBSYEEGNNNNNDINNINNINNINNINNINNINNSCNYGKAKKTTLLKRDIDKQEGY 177
QY 496 -NNINNNNNNNNNNNNNNNNNHLLPQPNYAFDTSDPSLDDMGNCHLNNDKSIFLHK 554
D 178 NNEITTLNNKNNKNN 205
QY 555 NLRDNIKLKRRSSCDNINMKKNNLHLARASVSKLTFMSYDPOKKNKTFPKSNIKME 614
D 206 NNNKNN 247
QY 615 DNTPKDILYSRVSN--NNGNVLLGINKTHTDSTKDNHNDK--INGVINIINN 669
D 248 KNYOKT--EQNNLNNHTYLNNTNNIINN-----NNGDNQYAYINN-FYHIYHN 250
QY 670 SNNNSINNSMNSGINNSNNSNSIYKSN-----YNSNOSISDVOIRYNNEM-DPSNKNND 724
D 291 NNTNHHYRQNNIIPICNINNAFNIETKLNPPYHDN-----HAYTNVYSTONK-- 341
QY 725 IFPDALSCDNNM-----YDNTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 775
D 342 -NNMTKQIGHYGINNEDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 389
QY 776 YAVNSHLFOPDNNKNTSNIQNTNTKKNODGAVNTSMFCHYLANDKYLID-LNKE- 833
D 390 POLDNYKTKNNENQGTNNENQGTNNENKCTNNFNNAKNIKENINTNKIEHLNHSI 449
QY 834 -QKDKIHOCDDNIIIONRDFEKKKKTTFYNNNNIYIYANNNGNNNSPRMKYGLCG 888
D 450 YNFVYPPNKIUYANGLYNNNNIYSTQKNNNNINFNHIESPINOQHNTTFV--N 504
QY 889 SHTSIDMKNNKNNEMKNNEMKNNHFKSNSSSSSSNNNN--IYNNINDDFOND 946
D 505 NDNINFPREPNTKKK--KKEKKKNIHNNNNNN-----NNKKCYKDIIN--QND 549
QY 947 YCHNDNTFTIRKKNNTINSNIYQNDIITYINSLSNDYMSNTLLHFEKXYTPTLSTNED 1006
D 550 H-----NNSIINTN--QNFHDINN--KTEQNLQKKKKNKKSQVSKOS 588
QY 1007 IYKMEGKIRLDODDKYDNDNNVNDNNKANNVNDV--DNNVNDVNDV----- 1056
D 589 NNNKNNKNSHLK--KQININTNNNNNDKNSHISKVIVDNDKLSHADSNELVTK 644
QY 1057 -NDKANNVDN--NNVDDDDVDVDFHNKKNFNNNEY-----LSYF 1091
D 645 GKKKKNTKKKKKINNINNVANNVNNNNINSMNNIISNNVYNNNNNNPFPYNNVNIQKD 704
QY 1092 QKRVDTIINNCLNSLDSM--YDTEKILNNILSKYAKKONVYIKYINEDIKMSL 1148
D 705 DSNIALLYNNKPN-IDENNQLNHNHMIQNNIMTNVNL--NNNLTTSFNVLNLYSY 762
QY 1149 E-----EIDKTAQ-SIYEKRYLLTLLLFKKNVDTQINNESDL 1188
D 763 EPEYENLMDLDYCARDISLYEK-----YDRGDLQONHKKYDI 802

RESULT 12
S71629
sensory transduction histidine kinase dhka - slime mold (Dictyostelium discoideum)
C/Species: Dictyostelium discoideum
C/Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 24-Sep-1998

C/Accession: S71629
R/Mang, N.; Shaulsky, G.; Escalante, R.; Loomis, W.F.
EMBO J. 15, 3890-3898, 1996
A/Title: A two-component histidine kinase gene that functions in Dictyostelium developm
A/Reference number: S71629; PMID:96324397; PMID:8670894
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-2150 <MAN>
A/Cross-references: EMBL:U42597
A/Experimental source: strain Ax4
C/Genetics:
A:Gene: dhka
A:Map position: 6
C/Superfamily: response regulator homology
C/Keywords: autophosphorylation; phosphoprotein; phosphotransferase; two-component regu
F/2027/2142/Domain: response regulator homology <RRH>
F/2076/Binding site: phosphate (asp) (covalent) #status predicted

Query Match 9.9%; Score 676; DB 2; Length 2150;
Best Local Similarity 24.5%; Pred. No. 6.8e-21;
Matches 242; Conservative 146; Mismatches 255; Indels 344; Gaps 41;
QY 416 TFGKYSLSLNDIDKIKKKKKGLINYKSTLYNDNTINKKNNNNNNNNNNNNNNNNNN 475
D 85 SYGNHSFHHVSPSP-----SYDINNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 137
QY 476 N-----NDSSSM--VNNMIH 489
D 138 NKNVNNNNYYSPINENSISKLSIESVLNQPHNFNLSSNNNYLNNSSLSLHINQSVNS 197
QY 490 MINNNINNN--INVNNNNNNNNNNNNNNNNNNHLLPQPNYAFDTSDPSLDDMGNCHLNN 545
D 198 LSNNNNNQNTQPIKNN 237
QY 546 DKSIFLAKKLNLDN-IKLKNRSSCDNINMKKNNLHLARASVSKL--TFMSYDPOKKN 601
D 238 GN-----NNNNITDPSFKRSHSTYETNIGS-----HQRRKSIQSLIANGAHSFKLK 288
QY 602 DNTFFKSNINKKEDNTPKDIYESRVSNMNGNVLLGLNKTHHDISTDKNHNDKINN 661
D 289 -PLSSSTPSTV-----NTCGAVNNNSN-----NNNNNNNST 319
QY 662 GVNIIN--NSVNSI-----NNS-----NNNSINNS-NM 688
D 320 GSIGALPMRSPDGNITITTESTGANNSPRNSGSCNGGIPLSPRNLSLSNGVNV 379
QY 689 NSNSIYKSNYNSQOS-ISDVOIRYNN-EMDTSKNNNDNIFPDALSCDNNMYPNTNNNN 746
D 380 SPENHILNNLNNNSNLPLSPRHINFIHIVNSLNNNN-----NNNINPNNNNNNS 431
QY 747 NNNNN--NNNNNDIVENYNNRBDGTNNSMKLYAVNSHLFOPDNNKNTSNIQNTANK 803
D 432 NSNNNNPNNNNHNTSPRSNTSPRSN-----GSGTTISPRNTSNNNNIIN----- 478
QY 804 NODGAVNYSWNFCHVYLANDKYLIDLNKKEOKDKIHGCD-----NNTIIONENDE 854
D 479 -NINNNTILTPRNSPRLE-NVNPNTNSRLATSLNSTLPIYSSLT 522
QY 855 KKKKTFYNNNNIYIYANNNGNNNSPRMKYGLGSHTSIDMKNNEMKNNEMKNN 914
D 523 SSNNNNQSNNTNPISINNNGRN-----GHCIOIISBEILGNKRVVYVNNNGNN----- 570
QY 915 HIKSNNNSSSSSSNNNINNNINDDTQNDYCHANDTFTIRKKNNTINSNIYO----- 970
D 571 -NNNNNTNNTSNTNNNTTNNNN--NNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 607
QY 971 -NDIITYINSLSNDYMSNTLLHFEKXYTPTLSTNEDIYKMEGKIRLDODDKYD 1026
D 608 RTKGHS--KTNSLDFETS-----SNNGGDSISGAGSGSLARRKNDND 652
QY 1027 DND-----NNVNDNNKNNVNDVNDNNVNDNNVNDNNVNDNNVNDNNVNDND 1073


```

Db      837 QMTNQ-NANQWK-NQVNSHINSQVAGHMANMNFNGNNSGTNMLGNMMLNNVNSMDN 894
Qy      808 NV--NYSMNF---CHYMLNDKNTYLIDLNKEOKDKNIHGCDNNIIONRNFPEKK---KKT 859
      895 NVNNNYSNNYPCSGNNGNNTSCVMNNNGAEIDMD---DNVIEH-NQHEHNNMLEKE 949
Qy      860 NFYNNNNIV---IVNNMGNNSPRMKYGLCSHTSIDMKKNEMKNEMKDNKDN 915
      950 DSINNDSIITTTTATNINSNNNS-----NNANSLNMLVMSNGNGN-TKLTNI 997
Qy      916 IKSNNNNSSSSSSNNNNIYNNINDDPTFQNDYCHNDVTFTRKNNNNI---NSNIYON 971
      998 INNNTNNNNNNSSSGGVNNTNN-----VNNNNMNTYNNNNNNIINDGSSVYGN 1049
Qy      972 DDIITYNSLNDYMSNTLLHFKEKYTYPTLSTNEDIYKMEGKHRLDPODKDDNDN 1031
      1050 ESMTHKKQVNV-----CINNDE-----DDLGNMNNHNN 1079
Qy      1032 NVDDNNKQVNDVNNNDVNNVNDNDNKN-NVDNNNVNDDDVDVFN-NIKNFNNNEYL 1089
      1080 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1139
Qy      1090 YFOKNVDTIINNCLNSDISMYDTKEIILNILLSKYAKKQNVYIKYINEDIKMSLE 1149
      1140 EGRIN---FTHAGKHVDKNNMOSGTTVVNVL-----NNDENYIKSNINNTYND 1186
Qy      1150 EIDTKASIEYERKVLTLKLLLFKQVDTQJNNETSDLRKDLVYCHICNNPD 1203
      1187 NMKNNDPIDDDNNNIVENMLRVEKNI---LNNNDE-----HINNTEED 1229

```

RESULT 15

118429
 hypothetical protein C0345w - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C:Accession: T18429
 R:Lawson, D.; Bowman, S.; Barrell, B.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: Z18935
 A:Accession: T18429
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1711 <LAW>
 A:Cross-references: EMBL:Z98547; NID:e1325376; PID:e1325381; PIDN:CAB1106.1
 C:Genetics:
 A:Note: PFC0345w

```

Query Match      8.9%; Score 611; DB 2; Length 1711;
Best Local Similarity 23.8%; Pred. No. 2.0e-18;
Matches 284; Conservative 202; Mismatches 406; Indels 300; Gaps 57;

Qy      127 IFFPLNYDKDNTKGIISFEYIYCGLYDLQKRVNVALENGKEVYVVKDKILRYLT 186
      665 ILPFTDY-----MNEIIS-YQIKC-----KRSLSHKSWEKRVASRSL-YLCIHT 709
Db      187 KEELIMIDGVLLRKIGVNSQNDSESRSHAILNIDDKINKTSIGKIFIDLASERG 246
      710 HSKYI-----KSNKYQND-----FDENVSHKNDVAGSIAY-EMEQVE-- 747
Qy      247 ADVTSQNKQOTDGANINRSLALKEC-IRAMDKNKHIPIRDSLETKVLADIIVGSKS 305
      748 -----INEERRDQ-----EMLVGDEMEENKNIENYHIDMASSEISN-KENCLIQNMS 796
Qy      306 IMTANISPTISCEQITLNTLRYSRVNFKKSTCINEEDDTTERSIID-----SKGSE 361
      797 NNSINDIESITICSNV-----EIKMF-----NDPNTITITILEKSNYSYIE 838
Db      362 MNASSIENVYKSNHLLNNNNN-----KIRGKINDK 394
      839 LTRNDKDLV-GWLLSQNNKDMKELIYILCAIKKGNPKARIPFYKINNNNSMFM 897
Qy      395 IERNN---ILKNKSFDPREGF---TSTF-----GKY--SSINDIDIKKKNKKGLINY 440

```

```

Db      898 YHENNISTTYKNSANLIEPSYFINTSEHEKDERDQKYLEASINDYMSDKKKR-----Y 953
Qy      441 KS-----LYNDNTIKKNNN-----NNNNNDNNNNNDNNNNNNNDSSSMVNNMI 487
      954 DSIESLRGSDKKNDDIYQGGHSSLLYYDDNNDDNNNNNNMYDSSSHNNYIILTNDR 1013
Qy      488 NHHNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 547
      1014 LNN-DNFINNNLEINNQQKRIEKLLEYLN-----NVKLTKSINYSQSNNTN---SKDH 1064
Qy      548 STPLAKKNIJDNIKKNRSSCD-----NIMKKKNLHLARHVSGLTINPSYDQKN 600
      1065 NISDSKKKEDTLNLSRKSSEYNNKILOSTSNKSLNGAY-----ENNLFSGKKKK 1116
Qy      601 KNTFFK--SNTKMEDNPKDILYSRVSMNGVLLGKN-----THHDSIKDEN 653
      1117 K-GTVLKDIHINDIQDKPEDL-----NINCNTKVIENEERHLLPLEEYVLSSDEK 1170
Qy      654 HNDKINNGVINIINNSVNSINSNMSINNSNMSNYSKYSNV---SNQSLSDVOI 709
      1171 FGLNKIK-----NDNNIYKQHQYHNLXDQNGKHILFPTNKQVSLQRNNNINSYIK 1223
Qy      710 RYVNEKDTSNK--NNDN-----IFPDALSCDNNMYPN---ITNNNNNNNNNNNNNN 755
      1224 TNYHEVEKNNKQQRNYDFTCDKKKIYYNIINSDKDIYHNNIITYKNEKEGIGITHLR 1283
Qy      756 NNIDVENNN--NRDGNNSMKIYA---YNSHLLFQPDNNKNTSNQNTNKNQDGVNV 810
      1284 NDKDITNELLKLDGKEFLDTFKDSYIDCHN-----KKNILNM-TNNKKEHQI-- 1333
Qy      811 YSNMPCYMLNDKNTYLIDLNKEOKDKNIHGCDNNIIONRNFPEKKKTTFYNNNNIV 870
      1334 --IDVADKIFNETNMTMNNKIYDDKVNH-----EKKCTINDVIHNNMDL 1378
Qy      871 NNNMGNNSPRMKYGLCSHTSIDN-MKNNEK-----KNEMKDNEMKNHISKNNNS 923
      1379 STSIKNN-----ENLFIDTYQKNRIGDIYMRINILQEDDDDDNNNNNNNNNN 1428
Qy      924 SSS-----SSNNIYNNINDDPTFO--NVCINDNFTTRKKNNTNINSNIYQNDI 974
      1429 NNNKILFEYTYKDQMLHNNKNNLEGTSEFSDPIERKNKIKIKNNES-----YKIDE 1482
Qy      975 IYINSLNDYMSNTLLHFKEKYTYPTLSTNEDIYKMEGKHRLDPODKYDNDNNNDV 1034
      1483 SLLSNEKNNKVSLLIN-----NKKDSSVVMKNNKNNKNNKNNKNNNNNE 1526
Qy      1035 NNNKNNVNDVNNNVNDVNNNDKNNVNDNNVDDDDVDVFNIKNFNNN-----E 1086
      1527 NNNKNNKNNNNDSFSKDNLLINNDDNNNNNNNDSPSKDNNLLINNDDNNNNNNNNKYIKE 1586
Qy      1087 YLSYQKN-----VDIINNCLNSLSDSSYVD-TTEILNILLSKYAKK 1131
      1587 IIDDEKDKDIHRDNIYIKDVSPLINHPN---INSRKQRTIEPL-KIINGKKNKLIK 1642
Db      1132 D-NVIKKYINEDIKNNLSLEIDTKASIEYERKVLTLKLLLF--KQAVDTQ 1180
      1643 DLKTIQOYERKIRKTKTIQMDQENKPPPSKKKIKNNKSLINDIDDDQNVDSQ 1694

```

RESULT 16

118477
 hypothetical protein C0495w - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2002
 C:Accession: T18477
 R:Lawson, D.; Bowman, S.; Barrell, B.
 submitted to the EMBL Data Library, November 1998
 A:Reference number: Z18937
 A:Accession: T18477
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2523 <LAW>


```

Db      1655 LNKKDSQVDFNNFIISN---MYNNSDFKKKNIGNVLQYLDYDQVYKKNVHIFNM 1711
Qy      64 TVDNFTEYENTIKPLIDLYENGVCSCFAYGQSGKTYTMLGSGYQSDTRGIFQYA 123
Db      1712 YISNVTHKE-----MENQ-----MYRKEGNR----- 1734
Qy      124 AGDIFPELNYDKDNTKGFIFISFEYICGLYDLQCRK-----MYALEN-----GK 171
Db      1735 ENNTYDKINIDKKNFG-----QPIY--KYVDLTSKNTLWDSJNGCJESSYVTKGSN 1786
Qy      172 KEVVVQDLIL--RVLTKEBILKMDGVLLRKIGVNSQVDSRSHAIILDKDINKA 229
Db      1787 TKIRIDIEKIGKMLDTSNQLVMKSLFLEFEDKLG-----SYLLQNDMMFSTTKN 1836
Qy      230 TS-----LGIAPIDLAGERGADRVYQKQOTGDANI--NR 265
Db      1837 FKOBISYIQLSLYEKNFQVLFLSNIYYSDVANKKIGAFSLFGKSNVHIAENIRNT 1896
Qy      266 SLALKECIRAMSDK--NHI--PRDSLTLYLVDIFVGSKSIMIANISPTISCEQTLN 323
Db      1897 LILSMGLCMKKNHSHKFNHLNYIREAVYSKEIFHIYEKTI-----YLN 1939
Qy      324 TLYSSSVKQF---KXKSTCINEBDTNTERTISILSKSGSEMMAS-----SIENVV 371
Db      1940 IAKENEFENILFDKDSVGIIPTEIIN---KIGNKQKYKTYNVDQINLYIKIKNLL 1995
Qy      372 IKSNHLSNNNNKINRGKINDKIERNNILKNSFDEREGFSTPEKYSIADIDIKIK 431
Db      1996 LSSIH--ISNSNTSDLN--KIN-----VEGQRYEDP-----VKR 2025
Qy      432 NKKGGLINYKSTLYNDTINKKINNNNNNNNDNNNDNNNNNNNDSSSMVANNIMHI 491
Db      2026 TLSIMRITNNSIRNEMDNKNNKNNNNKNNNNKNNNNNDNNNNQSN-----YHS 2078
Qy      492 NNNINNNINNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 548
Db      2079 HDQIKGVGVYKNNSSSTNDN-----KIDLSNMENNNNNNNIMVMS 2117
Qy      549 IFPHKKNLRNLIKLRSSC--DNIMKKKNL--HIAHVSGLTYMF---SYDPQKNKD 602
Db      2118 DYIINNGFTSNGTCSBPNCNDLMSQKEEIAHIYKSKELHFFLLQLVDPIL-- 2174
Qy      603 NTFEKSINIMKEDNTPKDLYESRNVSNMNGVNLG--LNKNTHDISTKDENNNNKIN 660
Db      2175 GSIYQEKESVLOSSLKSLPMISKRLKIGIILLENDKYNNTKEDLIKNNRLNDESK 2234
Qy      661 N-----GVINIINNSVNSI-----NNSMNSINNSN-----MNSNSIYK 695
Db      2235 NMEFKYVSCBILYIDINNRKICILKODICEDYLKENLSNVAGSCLCYKFIISLYYL 2294
Qy      696 SN-----YNSGOSIDVQIR-----YVNE-----MDPSNRY 721
Db      2295 SNIDNLSFPIHNTTDDISNVLSLEKFLVYKFRGIIINKESEYFKEIILKYGEDQYVN 2354
Qy      722 NDNIFFDALSCDNNMYENITNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 781
Db      2355 DDYMSINNTKNNNNHNSNNHHKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2403
Qy      782 NLFQPDNNK--NTSNIONINTNK-----NODGVNYSMNFCHVNLDKNYLIDLANKK 835
Db      2404 -----NNKGNINIKVHKTIRKQIDRLHQLNSNINLEKEFYIKL-----IYEL--DKKM 2452
Qy      836 DKNIHGDNNIIONRDNFEKKKKKTTFYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 894
Db      2453 IKRLANSSNNFI--RNDY-----FSELLIOHNGSNNEETKYKNDYHHPVIDI 2500
Qy      895 NKNKNEMKNNEMKO-----NEMKDNHFKSNNN-- 921
Db      2501 NKESYELLHSLKTIYNSLFIIRHPLNFQYFIIVDVULGLFKIVNEKRIYQFQK 2560
Qy      922 ----SSSSSSNN-----NIYNNINDDTFQNDYCHNDTFTIRKKNNTNINSNTYQND 972

```

```

Db      2561 EGCCKESGHEHNSKGINIVSKGEENMG-----KRRNKRTKYQVQEEK 2608
Qy      973 DIITYINSNDVMSNTLHPKKEYTPTLSTNEDIYKMEGKIHLLDDODKYDDNNNN 1032
Db      2609 E-----NVTL-----DWTYKTIKNDIDIY-----DEHKCDNNNDN 2641
Qy      1033 VDNNNKNNVNDVNNNNVNDNNVNDNDK-----NNVDNNNNVND-----DDVDPEFNK 1080
Db      2642 -DNNDNN--NNNDNNNDNSNNSNNKLGIVYVTHNNILKDNMGKAAFKCDNDIYHSNK 2659
Qy      1081 -----NPNNEYSYFQKNDTIIN-----CLN-----SLDIS 1109
Db      2700 KSTQKNNRKKSSYIKLKKYKNSSPFTNNIFDLNRYDKISISCLNLIISIVRLVDIN 2759
Qy      1110 SMYDPTKEILNNILSKYAKDNVTKTYNE-----DIKMSLEEDKTAQSIYK 1161
Db      2760 SFF-----FLSK-----IKIYIHDCISTNISYVFLFRIT--LOWVNS 2798
Qy      1162 RYVLTLLKLL-----FKQNVDTQ--INNETSRLKDLVWCHI--CNNPDDQFHF 1208
Db      2799 RSIKMRVRYKVSSEEEYCDKGSFYNSIEREMWKEETKEVGEQYKEDILNDRQNNNTSL 2858
Qy      1209 YAY--SRLEKDIINLMRQWCESENRLLYQFLV---VEYQKNSAVYLVNSSNN-- 1261
Db      2859 LSYNSNFER--YFSLF-----SPNSTVSYGTLFDLKQEQKLEKSNVYIERNNMD 2910
Qy      1262 -----GDILLNKLYQD-----NIKMSDHNH 1286
Db      2911 KATDLYDNYVSDWTKENRTLGHENKGTPEKYNNNNNNNNNNNNIN 2955

```

```

RESULT 18
A23770
Asparagine-rich protein - malaria parasite (Plasmodium falciparum)
C/Species: Plasmodium falciparum
C/Date: 21-May-1988 #sequence_revision 21-May-1988 #ext_change 09-Jun-2000
C/Accession: A23770
R/Stahl, H.D.; Bianco, A.E.; Crewther, P.E.; Burkot, T.; Coppel, R.L.; Brown, G.V.; Andl
Nucleic Acids Res. 14, 3089-3102, 1986
A/Title: An asparagine-rich protein from blood stages of Plasmodium falciparum shares d
A/Reference number: A23770; M0ID:86176787; PMID:2421257
A/Accession: A23770
A/Molecule type: mRNA
A/Residues: 1-537 <STA>
A/Cross-references: GB:M24328; GB:X03716; NID:g160091; PID:g160092

Query Match      8.7%; Score 597; DB 2; Length 537;
Best Local Similarity 28.5%; Pred. No. 3, 1e-18;
Matches 212; Conservative 92; Mismatches 193; Indels 246; Gaps 39;

Qy      424 NIDIKIKKKKGLINYKSTLYNDTINKKNNNNNNNNNNNDNNNDNNNNNNNNNDSSSV 483
Db      2 NNNNNKNNNDGNTINYQNT--NEFKDKK--NMFKQYNNNYKFDENNNNSNTVHSKN 57
Qy      484 NNMIMHNNNNIN--NNINYN-----NNNNNNNNNSHNNHLPQPNYAFDTDS 530
Db      58 SIVEEHLRNSIDMNSNNTNTNOQTRFSFMEHENEKNTYTG----- 104
Qy      531 DFSSLDNNKCHLNNDKSIFLHKKULRNIKLRSSCDNINKKKNLHLARHVSGL 590
Db      105 -----GNANN-----IHFKNKYD-----NNSSMKNTDNNK----- 130
Qy      591 TMFSYDPQKNKNTFFKSNIMKEDNTPKDLYESRNVSNMNGV-----LLGLKNTHH 645
Db      131 TDTSTN-----MKGTIN--NDNNNDVIRNINNTINEYGSAKNKFYTYNMKNKLNK 179
Qy      646 DISTKDENNDKINNNGVININNSNVSNNSNNNSNN--SNGI--YKSNVNSQS 703
Db      180 FTQNNNDNNINIEDN-----NNNNNNNNNGVSNTQNNNNKNNNSINTIRNLNNNN 233
Qy      704 ISDVOIRYVENEDTSKNNNDNIFPDALSCDNNMYENITNNNNNNNNNNNNNNNNNN 763
Db      234 INN-----MMKMGSGODKQNS-----NNNFYNNVYQNRKNSMNNNNNNNN 274

```


Db 1270 IINDVLKSDKFKEMINSKE--INIKDFCVKDNNDICLN-LNELNGQY-----NSNT 1320
Qy 1203 DDQ---FFAYASRLKEDIILIML-RQWCESENRL-----YOFLAVEYQKSA 1250
Db 1321 TEKIDIDHDFDYE-KNDVADIKIHEDREYCDSEKLYLPAADIKKRSIFIKINKSKRE 1379
Qy 1251 NSVLLANVSNGDIILKCKLVQDNINKSMDN-NIH 1286
Db 1380 N-FLIISYMEIEYFLIKLKVIEIPSIILNKGMH 1414

RESULT 20
H71621
serine/threonine-specific protein kinase (EC 2.7.1.-) PF80150C - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: H71621
R:Gardner, M.J.; Tetteh, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MIMD:99021743; PMID:9804551
A:Accession: H71621
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2485 <GAR>
A:Cross-references: GB:AE001376; GB:AE001362; NID:G3845108; PID:AMC71820.1; PID:G384510
A:Experimental source: clone 307
C:Genetic8
A:Gene: PF80150C
C:Superfamily: malaria parasite serine/threonine-specific protein kinase PF80150C; protease
F:2087-2352/Domain: protein kinase homology <KIN>

Query Match 8.6%; Score 587; DB 1; Length 2485;
Best Local Similarity 20.5%; Pred. No. 4.1e-17; Index 616; Gaps 85;
Matches 356; Conservative 249; Mismatches 517;

Qy 21 KKSDIITVKNCTLYIDEPYKVDMTYERHEFVYDKVDDTVDNFTYENTIKPLIT 80
Db 359 KENININISKEC--HDEK---EEEKIIMVSNLVEE---KKOKKIMVSNLVLPI 407
Qy 81 D-LYENG-----CVSCFAYGQTSGKTYMTLSGQPYGOSDTPGIFQYAGDIPTPL 131
Db 408 DILLKNGHDEINKEICCKKKKSPFQNDIKSKMLYNNKSKSKSEKV----- 453
Qy 132 NITDKONTKGIFFSFEITYCGKLYDLQKRMVAALENGKEVAVKDLKTLVLYTEEL 190
Db 454 -LYTNKNKSNFTPIPIF--PLNKVGDKFKNSENIYDMYNNKKNYIHDKKIYTMYSNKKL 510
Qy 191 -----ILKMIIDGLV-----LRKIG 204
Db 511 QKHYTSTSNINLLYNNIGKVLNGLHLSSNMVCRINSNPYKSIILINNVPFYKRRKSN 570
Qy 205 VNSQND-----ESSRSHAILNIDLKINKNTSLGKIAFTDLAGEGADTVSONK 254
Db 571 SNNNNNNNNISSSSSSSSKKNHVIN-----KISSYNIHYKER-KDSFREN- 616
Qy 255 QOTDGNANIRSLALKECFRAMDSKNIIPRDSLETYKLRDIFPGKSKSIIMIA---N 310
Db 617 -----FLFFKE--KILPSKQDTCVFNERQ-----KDLFEKSNHEIKCVSSFPNN 657
Qy 311 ISPTISCEQT-----LNTLYSSRVKNF---KNSCTINEDDPTNTRISITLDS 357
Db 658 TSDDISHSVKNKEPFALKNNISIRHPIKENNIITYSGKS--FNHVQD--KERTVLKK 713
Qy 358 KSGEMNASSIENVYKSN---HLSSNNNNKINRGKINDKI-----ERNILKSKSPDKR 410
Db 714 KKEINDKNTFSSGCLINNINITYTLQNGVKNLMLGIRDSIYKIDENNNLKS----- 766
Qy 411 EGTSTFPGKSSLANDDKIKKKKKGLINYKSTLVNDN-----TINKGANNN----- 458
Db 767 -----CYNGNDSNNKKKKKKKCL-SFSCDIINDNITPYESDKKNSNNIKSMIDIF 817

Qy 459 -----NNNDNDNDNDNDNDNDNDSSMV--NNMINHMINNNIN----- 497
Db 818 NYKRSNLYNNLSSRSDSTVDMENKYNSEBYINIOPTKIYELSKRIANRYKLSMDEI 877
Qy 498 -NINVN-----NN 552
Db 878 FKVSLKEKKYIDNISNNMERVYKNEM--INERKISMDLILYPCDKRSLNNSCPVI--- 932
Qy 553 KKKLRKNI-KLNRSSCDIINKKK-----NHLAHSVSKLTMSYDQK 601
Db 933 ---IENNIRSEENKSSVILNKKKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNK 975
Qy 602 DNTFFKSNINKAEDNTPKDIYESRVSN-ANGVNLGLNKQTHDI--STKDEHNHN 657
Db 976 DNIYDGNINKKBEETTKDEYISREKQKYNKSKIRNDDYKQVLSYHTLDEBKKN 1035
Qy 658 KIN-----NGVINI-----NNSN----- 671
Db 1036 DMNNLIDMNNNEALIEFVNGVINNIILDRKDNNSRDKMEKEMEKEMEKEMEKEMEKE 1095
Qy 672 -----VNSINNSNNNSINNSNNNSIYKSN--YNSQSIISDVQIRVNE-- 714
Db 1096 MEKMEKEVEKELKQEMNNRNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1151
Qy 715 -----MDTS-NKNN-----DNIFDAI--SC-----DNNNYPNITNNNNNN 747
Db 1152 KAIYFPNYSVDYKNNNNNSININCKCDYNNILKEVDSCLAQKEENIFRPLFLNKCKD 1211
Qy 748 N-----NNNN-----NNNNIDVENYNNRD----- 767
Db 1212 KWKKRPNIKIKTIHHEEMKRIYQIINKAVPIYFNFRBYENPLNHLTYFPKNDLFX 1271
Qy 768 -----GTNNSMGLYAVNSH--NLFQPDN--NKN-----TSNQN----- 797
Db 1272 LSKYKSMNIRNLVLANKHINNYYDMKLYNQNTYTLKYQVANDNDHHICCKGGLDY 1331
Qy 798 INTN-----KNNQDGVNYSNFCYH-----LNDK-----NYLIDLNNEQKDN- 838
Db 1332 INNNISKECKAKD--KTYLAKIPIYKKKKARFINDIGSNDWYDIKKKYSNDENNY 1389
Qy 839 -----IHG-----CDNNIIONR-----NDPEKK 857
Db 1390 KLNEMKNISMSDEMIPTLSEHGNFPSCQPNLEKKSITYIDLNDVDSNMDP-TEE 1448
Qy 858 KTFYNNNNIYVANNKNNNSPRMKYGLCSHTSIDMKKNEMKONEKD----- 908
Db 1449 KNFVNNENDLP-----NTRKWKF--NFSKGNLFFNNKFPNVSGEDVPSFFPKM 1496
Qy 909 NEMKDNHISKNNNNSSSSSSNNNNIYNNINDQTFQNDYCHNDFTIRKKNNTINSNI 968
Db 1497 NLFRELKSNNSLKEISYKSNNSNNSN--KGDNNIGNNENNTVYTLASDEHISTKGDI 1555
Qy 969 Y-----QNDIITYTINSLNDYMSNTLHFREKYT-----YPTLSTNEDIY--- 1008
Db 1556 HSESFRDNDCTILKIGRSKYSIDITLYNEDKSNLENDFTINEYENCNIDVNEE 1615
Qy 1009 -----NKEGKHILDDODKYDDNDNNVNNNNKNNVNNVNNVNNVNNVNNVNN 1057
Db 1616 DRYNGTCSVGEKETEKNNEKNEKNEKNEKNEKNEKNEKNEKNEKNEKNEKNEKNEK 1674
Qy 1058 DKNVNDNNVNDND--DVPFHHIKNPNNEVLYSTQKAVDTLINNCLSLDSSMYDPTK 1116
Db 1675 EENNEENNEENNEENNDIEKNDIKONNSQO---YKENI--IYNNNTNMDVDN--DDN 1726
Qy 1117 ELNNI-----LUSKYAE-----KDNVYKKYINDIKRMSD-----E 1149
Db 1727 NNNYVSTDEGIDIIKIKSEKNDYIYNDNIMIKINNSIDLNNIKNKNKNEPFLNYTEK 1786
Qy 1150 EIDKTAQSTY-----EKRVLLTKLLLFKNV-----DTQINNET 1185
Db 1787 DIHMSNSSYVNDKNNLFNNNEKTEKNTSLNDLYRKKBELDDEKISEYKDTNLTNNT 1846


```

Db      1308 IKGRVYKINNLIVIKRELFIYH-KCLVIMKLEKKKKKKYDNGKN---DIDINTICM 1363
Qy      185 LTKBELIKMIDVLLRKIGVNSQNDSSSHAL--NIDLKINKTSLGKIAFIDLAG 242
Db      1364 IDONKLL-----DNEKIKMELIDITVANKKDCSN-----WVD 1397
Qy      243 SERGADTVSONKOTOTDGANINRSLALKECIRAMSDKNHIFPRSELTGVLRFVGR 302
Db      1398 DEKNEDOTSNNTTISNG-KINNVVLKKNYIILND-----ECKL----- 1437
Qy      303 SKSIMANISPTISCEQTLNTRYSRVKNFKKSTCIN-----E 343
Db      1438 -----NIDIFSCITKSCTFPDSNELYYLEKILCSNFVVRNSHISIDIVYSN 1490
Qy      344 EDDNTTERISILDSKSEMASSIENVIVISKHLS-----NNNNKINGKINDIKER- 397
Db      1491 KDIITCSISLNFNNYNNNNNNNNIKNELPKSNEMIQKEILNNNNHTEPKKENDMTTPF 1550
Qy      398 ---NNILKNK-----SPDKREGFTSTFGKYSLLNDIDIKKKNKKGLINYKSTLYNDT 449
Db      1551 STGNDILKILLPDISGNNPDIICINIFIK-----EKNVLYLEKKSITNLK 1600
Qy      450 I---NKKHNNNNNNNNNNNNNNNNNNNNSSVNNMIMMIN-NINNNINNVNN 504
Db      1601 LILCTEYFHHFYILINKTEHRLKTFILNIMPMLFDIQENLFTINLICDOLYINN 1660
Qy      505 NNNNNNNNNHNNHLPQNYAFDITDPSLDMNCILNNDKSIPLHKKRLADNIKLN 564
Db      1661 NNNNNNNNNLCLILHPEY---YSQIIIPYLTCKNNKNVNDVSFKK---NFFKK 1713
Qy      565 RSSCDNIMKKKN---LHILAR-HSVG---SKLT---MFSYDPOK---NKNTTF 606
Db      1714 FKVDHLRKQYNGSFLNIQDHSKGLYERDKITRDLLSSFSQYTSILANQNVF- 1772
Qy      607 KSNINIKMEDTTPDI-LYESRNVSN-NGNVLLGLNKNTHDISKDEHNDRKINNGVI 664
Db      1773 ---IDNND---BDLFIYFLANCILNIPNI-----TNNVNS---QYND-KISD-L 1813
Qy      665 NIINNS------VNSINN-----SNNGSI---NNSMNS 690
Db      1814 NKTKNKKEERSFWSSSLKLLTKFDEEIFSTNDLRYEEYEMFISNLYILKKNKTISS 1873
Qy      691 NSIYKSN-----YNSNOSIDYQI-----RYVEMDTSNKND---NIF-----F 727
Db      1874 EVFFPSEYFLPYLNLPEFLRTGLWLTLLGLKADYDTINLRANDRNICVQYITAKK 1933
Qy      728 DAISCDNNMYPNITNNNNNNNN-----NNNNNNNNIDV-----ENYV 764
Db      1934 NSKSWDQNPPEVYNNMTDENNTTTTKKKDDNDQNDIYIHLIMNIVNRTLYTERLN 1993
Qy      765 NRGCTNSMKLYA---YNSHNLFOPNKNTSNIQINNTKKNODGVNYSNPFCHYNLN 821
Db      1994 NDRKINNKKKYEFTFTNMDIDIFNDN--NCINIIVNEDKKEE-----NIK 2037
Qy      822 DKNYLIDLNKBEOKDY--IHGCDNNITQNRNDEPKKKTTFYNNNNIVIVNNMGNNNS 879
Db      2038 DLTKKLTKEGKGVNDEFTQVTDNNI---EINPKKSTQNEEO---PNTINTINENG 2090
Qy      880 PRMKYGLCGSHTSIDNNKNNEMK-----NNEMKDNEMKDNHISNNSSSSSSNN 931
Db      2091 NMVYSILNSTLTLNNIHLKRWKYLINTYCFNNYIMFFQTQKYLRLRIKKAFLRS 2150
Qy      932 NIANNINDDTF-----QNDYCHNDTFTIRKKNNTNINSNY-----QNDIITYIN 979
Db      2151 LKEDDFDIDIKSYKRYKEINYC--DSNY---KNNKADTAQOYTHLKEDEKELINFD 2204
Qy      980 SLNDYSNNTL---LHFKEKYTPTLSTNEDIYNKMEGKIRLDODKY----- 1025
Db      2205 HIINYSNNIWSQIHF-----LTHGISLY---FENVGEVDEEDAMFYCLQKKKI 2253
Qy      1026 --DD---NDNNVYDNNKKNVDN--NVNND--NVNNDVNDKNNVNDNDVDDVDFH 1077

```

```

Db      2254 NEDDNKNGDNKNGDNKNGDNKNGDNKNGDNKNGDNKNGDNKNGDNKNGDNKNGD 2313
Qy      1078 -NIKFNENNEYLVSQKQVNDITINNCLNSLDISMYDTR-----EILNNILSKY 1127
Db      2314 GNNKNGDDN-----KNCDD-MKNC---DNKNCDDNKINSDEGVCEBSRKLLIY 2361
Qy      1128 KAEKQNVK-----KYINEDIKMSLEIDK---TAOSIYKRV-----L 1165
Db      2362 ---KNIIKENCEINSLNIYKYIRIWLRLKLOVDVFCILNEVINDKKKICFSSL 2418
Qy      1166 LTKLLILFKQVNDIQNETSLD-----RKDLVWCHICNNPDDQFHYAYS 1212
Db      2419 VNSVLSYVEHDFWLKNTDVLCSFJETOPTLREKQVINDINNEKSDPLFKYKKN 2478
Qy      1213 R-----LEKDIINILMRQIWCESENLILYQFLVEYQNSANSVLLNVSNNDDIT- 1265
Db      2479 KKEIKRYNDKFEIIT-IPFIKEKEEENLMDNLIKI-NQNDIKIDYNDMANSMDIIT 2536
Qy      1266 -----LNNKLVQDNIRKNSMDHNI 1285
Db      2537 NEDDKYLSNYINNNT--KNIKKIKNNNDV 2564

```

RESULT 23

```

T18402
asparagine/aspartate rich protein - malaria parasite (Plasmodium falciparum) (fragments)
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 09-Jun-2000
C:Accession: T18402
C:Barale, J.C.; Candelle, D.; Atchal-Bonnefoy, G.; Dehoux, P.; Bonnefoy, S.; Ridley, R.;
  Infect. Immun. 65, 3003-3010, 1997
A:Title: Plasmodium falciparum AAP1, a giant protein containing repeated motifs rich in
A:Reference number: T18929; MUID:97378065; PMID:9234746
A:Accession: T18402
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3844 <BAR>
A:Cross-references: EMBL:Y08926; NID:e1154302; PID:e111435; PIDN:CAA70130.1
C:Genetic8:
A:Gene: aarp1

```

```

Query Match      8.5%; Score 581; DB 2; Length 3844;
Best Local Similarity 20.7%; Pred. No. 1.2e-16;
Matches 315; Conservative 224; Mismatches 474; Indels 502; Gaps 63;

Qy      45 DMTKTYIERHF-IVDKVPDIDVNFYVNTIKPLIIDYENGVCSCRAYGOTSGKTY 103
Db      898 DMVVEIEIHIIKADKIFSDIYTKSYDSS----- 928
Qy      104 TMLGQPVGQSDPGIIFYAAGDIFPLNIYDKNTKGFISFVEIYCGKLYDLQKRM 163
Db      929 -----KIHIIYQFFYI-----IIRFNLHYIEYNIRPKTY 964
Qy      164 VALENGKEVVVKDLIRVLTKBELIKMIDGV-----LARRIGVNSQND 210
Db      965 EKYIERGKVA--VKKI-LIQLLASQDFQGEVNFPRKQFRHPSFVELVNNKYGYKTHMP 1021
Qy      211 ESRSHAILNTDLKINKTSLGKIAFIDLASSEGADTVSONKOTOTDGANINSLAL 270
Db      1022 RS-----NVALIKLNKMSWLYDAFWPL-----SAFKQFOT-----AN 1054
Qy      271 KECIRAMSDKNHIFPRSELTGVLRFVGRSKSIMANISPTISCEQTLNTRYSSR 330
Db      1055 EKCIKEENS-----YIGSR-----BEDKEYLNFSTR 1082
Qy      331 -YKN---FNKSTC-----INEDDNTERISILDSKSEMASSIENVIVYSKNH 376
Db      1083 KYQNLLEFYFKNSCLVSIILIVFIINDEFORKIEKKKELEKEBKMKKMKADVGEDDS 1142
Qy      377 LLS---NNNNKINRG-KINDKIERNNILKKNKSDKPREGSTSTFGKYS---LNDIDK 428
Db      1143 LNKQKGTNNNTCGVGVINEANRSS-----GMSQVVVYTESGINNNNDV 1190

```


Dp 560 EKKONKKEV--IMVDKNNDDTEKDQKKYDTSTY-SFNINQTLSTKYFYKNYVKRKMIGQH 617

Qy 837 KNHCCDNNITIQNRDEFEKKKKTNFYNANNIYIVNN-----NMGNNNSPR----- 881

Dp 618 NNHNHTNH--LHNTNNMH--NTHNIHNEKVLLDDTEKADPMNLGISFSPPAGILLPYH 672

Qy 882 -----MKYGCQSHTS-----IDNNKNNEMK----- 902

Dp 673 LGVSSLIEKIQTILMHHISIASSAGASTACCLSLGSLVNKCYPLENTISNVYHGCTOK 732

Qy 903 -----NMENKDNEMKHIKSN-----NN-----SSSSSS 928

Dp 733 LENTILNELNKYIVEDSYILANNRIGNVFVGITQLPYYKKLNNPFDDNDLSAILAS 792

Qy 929 SNNNIYNNINDDFONDYCHN-----DNFTT-----IRKRN 961

Dp 793 CNIPFYSSNIFVNFRRNCKTIDGFSTKKDPFGCPNTERIERIKVSPEPSDYVIGNKN 852

Qy 962 TNIN-----SNFYNDDIITYTINSINDVMSNTLAF-----K 993

Dp 853 SVISEPHLKVNHIELFCVKNIFHKYINNMIKEDYLFLLENLKDLERKIFPIYTFVK 912

Qy 994 EKTYTPILTSTEDIYK-----EMEGKHRLDDQDKYD-----NDNNNVDDNNKNV 1041

Dp 913 RYFTPT-LRKNETIDDKYESEEYEDEGESEEDDEDEEFYGANNNDODDEGDQKT 970

Qy 1042 DNNVD-----NNNVDDNVDDND--KNNVDDNNVNDN-----DDDD-----VDPHNI--- 1079

Dp 971 TNEKTKKKKKKNNNNNNNNI FNNNI FNNNI FNNNNNSCVGYSEKOFISTSIYASFANIKRQ 1030

Qy 1080 -----KNFNNNEYLS----- 1089

Dp 1031 MNEKIEKRKKEKKEKEXELGRKNNKCSKNNRRRYINKOSIHLMNLIRIKPKNLVYMN 1090

Qy 1090 -----YFOKAVDTIIINCLNSLDISSMTDTEKELLNILLSKYAEK-----DNVI 1135

Dp 1091 MWSEIELYLKINDIDFLQFNKHNYVONFNFYSITLIN-IMSXYSENFFAYANKIEIV 1148

Qy 1136 KKYINEDIKNNLSIEDITQAOSIYEKRKVLTTKLILLFKKVPDQINNETSDELKDLVMC 1195

Dp 1149 YKPL--LANNPFIETKQ-----YSKEDM-----NELDLVN--TYDMKYDKIE 1190

Qy 1196 HICNNN--PDDQFHFVAYSRLKDIINLIMLRQIWCESENIRLLLYOFL----- 1241

Dp 1191 FLKNNGYKLIDRY-IYFPYKLTIDI-LFFEKEIFLNDNMLKIDRKFLKXITIMIEVLK 1248

Qy 1242 ---VVYONKANSVL-LWSSANNGGDIIILAKULVQDNIXSM-----DHNN 1284

Dp 1249 EIFPEYVKRCITVIFPFAHKEHDV-MKKNYNNQYVNNNSMFMNTRGDHNN 1301

RESULT 25

G71609

hypothetical protein PF80650w - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000

A:Accession: G71609

R: Gardner, M.J.; Tetteh, H.; Garucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.; Partea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.

A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.

A:Reference number: A71600; MUID:99021743; PMID:9804551

A:Accession: G71609

A:Status: Preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-2500 <GAR>

A:Cross-references: GB:A0001408; GB:A0001362; NID:g3845238; PIDN:AACT1919.1; PID:g384524

A:Experimental source: Clone 3D7

C:Gene: PF80650w

Best Local Similarity 24.1%; Pred. No. 1e-16;
Matches 248; Conservative 164; Mismatches 389; Indels 229; Gaps 52;

```

QY 365 LDSKSEMAASIEVYKSHLLSNNNNKXINGKI---NDKIERNNIJKNSFPDKR 410
Db 33 IDEPGLVSNNMSVSNISSTSTNNIGTNTMFPNNSKGFIIINPENNYKKNIC--TYLDHES 90
QY 411 EGFSTFGKYSALNDIDKIKKKKKGLINYKSTLYNDTINK-----KHNNNNNNND 463
Db 91 TNINGGVQYD--NHMDQMOMQOTQMOMQOMQOMQOTQMOMQOMQOMQOMQOTQMOMQOTQ 148
QY 444 NNDNNNDNNNNNNSSGMVNNMLNHMLN-----NINNNIYNVNNNNNNNNNN-- 512
Db 149 MNOQTOMQOTQMOMQOMQOMQOMQOMQOMQOMQOMQOMQOMQOMQOMQOMQOMQOMQOM 208
QY 513 --NSHNHLPOPNYAFTDTSFSSILDMNCHLNNDSIFLHKRLJDNILKRSQCDN 570
Db 209 VQOSGMNANPL-----IYDISEL--YNRKNEQKTF-----RD--EYSRTTIKA 251
QY 511 IMKKKNNLHLARHSVGSKLTMESYDPOKA-----KDNFFFSNINKME----- 614
Db 252 LINKTTPMIMNSVKNIEDTNSJNDENYVNCMSDEYTTNYISKRYENDOQVIYOG 311
QY 615 DNT-----PKDILYESRNVS-----NMNGVLLGLMKHTHIDISTKDEHNNDKINNGVNI 666
Db 312 NNTYPEENDNNIYKKNELISIFODBLKONI---VEINAYHD-----SRHKRPIDQVANY 362
QY 667 INNSVNSIINS--NNASIN-----SNKNSIYKSYNSQSISDVOIRYV--NEMDTSN 719
Db 363 INNYTNNNDNPPYRNSSTNNNGIAENNINVASF--NOYKEXQYDYLTFPTGIMERKN 421
QY 720 KNNNIIPEDAISCNNVYPTNTNNNNNNNNNNNNNID-----VENYNRD----- 767
Db 422 IMQOVUDYNEKINGNSI--NIQSSNQOAMQOVYNNNNVCMQOPIYNNHNKKNYMG 479
QY 768 ---GTNSSMKLYAVNSHNLFOPD-----NNKNTSNI--QNIINTK--NNODGVNYV 811
Db 480 SNYGANNMNVAINIGTNNNNVYNNNVYNNHMGTTNNGTNNMGNNNIGTNNMGNNNGNNIG 539
QY 812 SMNCHVLYNDKXVLIOLNKEQDKNIHGDNNIIONRNDFEKKKXTPEFNNV--NIV 868
Db 540 TNNNVYNNHMGNN--IGTNNVTHNNMGNNYGNNGYNGA-----NYGNNYMGNNY 587
QY 869 IVNNNNGNNNSPRMKYGLCSHTSIDMMKNEMKNEMKONEMKDNHIXSNN--NNSSS 925
Db 588 MYHNNMGSTNN-----MGNNYMGDNMBRNNNNMGTTNNMGTTNNMGTTNNMGTTNNMGTT 640
QY 926 SSSGNNNIYNNINDDTFOYDNYCHNDTPTIRKQNTNINSNYQDDII--YTINSIANDY 984
Db 641 NNMGTNNMGNN-----NMGNNTYIENDN-----MKNHHI-----DYIINYNNVNNVN 683
QY 985 MSNTLLH--PEKETYFTLSTNEDIYKMEBKHIRLDDQKYDNDNNNVVNNNK--NNV 1041
Db 684 MYNTMYNNVNNVNNVYNNVYNNVNNVNNV-----MNGNNMGNNHMLNHM 725
QY 1042 DNN--VDNNNVNDVNDND--KNVVDNNNVDDDDVDVFNHIKNFNNEYLTSYQKAVDTI 1098
Db 726 GNDRIGYMGNLNINSNNYNNVNNNSNTYGNNNNDV-----RNNNVN--YAGNNMYARCL 780
QY 1099 INNCNLNL-----DISMWDRIKELINMLLSKYAEKONVKKYIINEDIKMSL 1148
Db 781 NNNNNNNNTPOYIIPDNKKISAVHPFKETNYGTIILN--NBSQDYTLRSISGIVL--NFSM 837
QY 1149 ---EEDIKTAQSIYERKXVLLTYLTLLEFKXNDVTOJNNETSDDR--KDLVMCHICNN-- 1200
Db 838 IOSENDFSTPSNIEPVNNOPLANEPAIF-----AVNLELPEASELNDQVCDENSTND 892
QY 1201 -----NPDQFHYAASRLKXDIINLIMLRQICESENLRLYQFLVEXQNSAN 1251
Db 893 VIKSKPLVNDLDEHIIYS--EHLNNRILNDETLPSAOLNVEYL-----YGEHEYNQGRN 946
QY 1252 SVLLNVSSNN 1261

```


[illegible]

```

Db      2136 -----INVHNINNEKEIEFDDEHLNLCQGNKHIIQKKKKKKTVYS 2174
Qy      1170 LLLFKQVDTQINNETSLDRKDLVWCHICANNPDQDFHY-----AY 1211
      2175 ILVENSINLVYLEGVDFK-----HISNHVIVFVLGLIEARITINEIKKCEVAY 2228
Db      1212 SRLKEDITNIMLRQIWESSENRLVQLVVEYGNKRSANVILNVSSNGDIIILNKKY 1271
      2229 S-IDIDIHIMLMDIMAFGLDILGINRFGI---QKAROSTLMASFEE-----TNEHL 2278
Qy      1272 -VQDNKSKSMDHNNI 1285
      2279 FVSSFFKXVDEINNI 2293

RESULT 31
B71619
hypotheical protein PF80280w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession: B71619
C:Gardner, M.J.; Tellelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
  : Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
  : Science 282, 1126-1132, 1998
A:Title: Chromosome 2, sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: B71619
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2539 <GAR>
A:Cross-references: GB:AE001384; GB:AE001362; NID:93845139; PIDN:AACT71845.1; PID:938451
A:Experimental source: Clone 3D7
C:Genetics:
A:Gene: PF80280w

Query Match      8.1% Score 552; DB 2; Length 2539;
Best Local Similarity 20.9%; Pred. No. 1,2e-15;
Matches 322; Conservative 211; Mismatches 479; Indels 528; Gaps 68;

Qy      133 IYDKNTKGI-FISFYEIYCGKLYDLQKKVVAALENGKKEVVVKDLKIIRLTRELI 191
      820 IHIKNNHGIYVQKKEMY--QLYDNNNNNN-----NNKSD----- 855
Qy      192 LKMDGVLLRKIGVNSQND-----SSRSHAILNIDLKQINKTSIGKTAIFDLAG---SE 244
      856 -----ICLNRYNPNKCSSEKKTNNPNSSILKPK-KKKKKQMGKIVTNLVKDNKEE 908
Qy      245 RGADIVSQNKQVQTDGAN-----INRSLALKKECIRAMDSPKNIIPRDSLE 291
      909 EGNNTIITKNDOSASGKTNEHMQIRINDAETTONNTLHKENKLCCTKQDNKLIHTKINSKEN 968
Qy      292 TKVLR-----DI-----FVGKSKIMIANISPTISCEQTLNTLFYSSRVKNF 334
      969 EKKVKKYHHYHINNDGLVFLVGLFIKKKNTIS-----LKLNIINNVLVKYKNNI 1020
Db      335 KKKSTCINEEDDTERISILDSKG-----SEMMASSINNVYIKSHLSNNNNKI 386
      1021 YKIKTVMYQKDIYNYLLNIIILLVGVKIYIRQHNKLKNESEYNV-----NSQNI 1070
Qy      387 NGKINDIERNNILKNSFDRREGFTSTPEKYSLNDIDIKK--NKKGGLINVKSTL 444
      1071 GSKSKSKYVWFHTSE-----ISFNKKILRPFYKIQKKINK-----YKRI 1115
Qy      445 YNDNT-INKKHNNNNNNNDNDNNN-----DNNNNNN-----NDSSSVNMMI 487
      1116 MNQSAHINIKESKNIIISNVEKSVTSNYSNIISSNNISPYYSIKENNMOKKTNKCI 1175
Qy      488 NEMINN-----NINNINVINNNNNNN-----N 510
      1176 EHLNNYKIKYIYEKIIYKVTENNHHLSFKIVDAESFSDPFSICILFSHFILSNIN 1235
Qy      511 NN-----NSHNNHLPQ-----NYAFDTSDSSSLDDNCHLNN-----NDKSLF 550

```


Dh 1236 ENIIFKIKNHONIKESTRIYHVPLKLFPHNLFICTNNSTYTTKMLHPLONIQFY 1295
Qy 551 LHKHNLB-DNIIKLNRSSCDNINMKKKNLHLARHVSGLTWESYDPOKNDCTPEKSN 609
Dh 1296 RYKKNIFETNNQKIYNTYIHNKYEKIQNFPNNSKYVINDQSLYLYDTONDRHRIIFMST 1355
Qy 610 INKME-----DNTPKDI-----LYESSNVSNNM--- 632
Dh 1356 ILSLIFKNIIIPKCDNHSKFPLEFHYAKYLHIYONGSNOFIINTYVPODNNINILHC 1415
Qy 633 -----GNVILG-----LNK-----THHDSIK 650
Dh 1416 TKKKRPGRGSTPEKYGGEIKGNDIIKESDIKCDIIESDVVKNELIVERNIIIEK 1475
Qy 651 DE-----NENDKINGVINIINNSVNSINNSNNSINNSNSNS--TYK 695
Dh 1476 DEIKTDKTEPIKNDTSDAKSIKSTSVLSSESNLSOCNNKLTKEHMEKNVITK 1535
Qy 696 SNYNSOSISDVOIRYVNEKMTSKNNNDNI.FDAISCDNNMYPITNNNNNNNNNNNN 755
Dh 1536 NNNNDNN-----NENNENNENNND-----NENNENNNDNNNNNNNN 1570
Qy 756 NNIDY-----NENY 764
Dh 1571 NNEVYKPYKINGLONIINSCNIFCSKKNIKKIKKIKKIKKIKKIKKIKKIKK 1630
Qy 765 NR-----DGTNNSMKLYVYNSHNL.FOPDNKNTSNIQNTNTKNNQDGV--NYSNMF 816
Dh 1631 TRHNGTQINNKVLVINITPYILRYPNNSSKSLCTKIKKIKKIKKIKKIKKIKK 1690
Qy 817 HYNLDKXY-LID-LN-NKEQDKNIGHCDNNIIQNRN-----DPEKKTITFY----- 862
Dh 1691 IHVNNKGYKKIDITLVHKEIDTSKQHTDEKICKIKQKYLVDYVRKAYISLYMYNKK 1750
Qy 863 ---NNNNIYVNNMGNNSPRKMYGICGSHTSIDMKNNEMKNEKNDKNDKNS 919
Dh 1751 KGGDTNNKNIQKKKKEEKKKSYNLSKHSI--LNNRMKYNIIIDMKNNNFYK 1807
Qy 920 NNNSSSSSSNNNIYNNI-NDDPTFOND-----YCHNDN--TFTIR 958
Dh 1808 DN-----YKIIYYDELENDINISYLKQINILNTIIGMKNVGTPLSKA 1855
Qy 959 -KONTINSNIY-QNDIIYINSIAND-----YMSNTLHFEKTYTPTLST 1003
Dh 1856 IENNIIIDIDEYLKDEIKFDKLSIDFRYEVYTFISLYLAFYILTFPRNLSAPRDT 1915
Qy 1004 NEDITYKMEGKAIRLDDOKYDNDNNNNVNDNNKANNVN-NVDNNVNDNNNDNNKN- 1060
Dh 1916 GATIKHVDIIDEKINSKQNKQTEYDINDNNNNYNSDNHNLHNNKDKQHTSTYKQ 1975
Qy 1061 -----NVNDNNVNDDDVDVPHNIKFNPNNEFYSYPOKNDVTIINNCLNSL--- 1106
Dh 1976 KKVSPDVCETIYVDGPFFEN-----KAYDNDIIFYTYNKG-I-TFYNNKINDPCK 2024
Qy 1107 -----DISSM-----YDTKEIL-----NNILSKYAEKDNVYK 1138
Dh 2025 IAKKCIQEKONGEHOMNTVITVAGGIIEDKSEKVKLKLKNTILI---KRDIDEIYD 2081
Qy 1139 INEDIKMSLEBEDKTAQSIYEKVALTLKLLLPKKNV--DTQINN--ETSPRLQDLY 1194
Dh 2082 INDNIK-----PKNGNFKDIIHRRITLYDKLSNAFHIIPSENNINKYIHSSEYNYI- 2135
Qy 1195 CHICNNNP--DDOFHYAYSRLKDI--NLIMLRQWSESNLRLLYQFLVVEYONKSA 1250
Dh 2136 ---NNNELIYHSLFREFNYPFKQPLIGDITNYKIKDKNEKN-----DEKDEK 2181
Qy 1251 NSVLNVSSN--NGDIIILNKLVODNIKNSMDHNNIHK 1288
Dh 2182 NDEKNEKDEKQGD-----NNDNDNN--NNEDENNKKK 2215

RESULT 32
T28160

hypothetical protein - malaria parasite (Plasmodium falciparum)
C/Species: Plasmodium falciparum
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 01-Dec-2000
C/Accession: T28160; S23688
R/Kun, J.F.; Hibbs, A.R.; Saul, A.; McColl, D.J.; Coppel, R.L.; Anders, R.F.
Mol. Biochem. Parasitol. 85, 41-51, 1997
A/Title: A putative Plasmodium falciparum exported serine/threonine protein kinase.
A/Reference number: Z20482; MUID:97262159; PMID:9108547
A/Accession: T28160
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-2510 <KUN>
A/Cross-references: EMBL:U40232; NID:g1658332; PID:g1658333; PIDN:AAB54058.1
A/Experimental source: strain FCQ27/PNG
R/Kun, J.; Hesselbach, J.; Schreiber, M.; Scherf, A.; Gysin, J.; Mattei, D.; Pereira da
Res. Immunol. 142, 199-210, 1991
A/Title: Cloning and expression of genomic DNA sequences coding for putative erythrocyte
A/Reference number: S23684; MUID:91376328; PMID:1896607
A/Molecule type: DNA
A/Residues: 241, 'S', 243-244, 673-959, 'R', 961-977, 'S', 1493-1494 <KUN>
A/Cross-references: EMBL:X53019
C/Genetics:
A/Genes: FEET
A/Intons: 645/2
C/Keywords: surface antigen

Query Match 8.0%; Score 549; DB 2; Length 2510;
Best Local Similarity 20.3%; Pred. No. 1,6e-15;
Matches 355; Conservative 242; Mismatches 516; Indels 638; Gaps 81;

Qy 2 NSKIKVAVKRLPELEKEDKSDIITVGNCTL-----YIDEPYK-----V 44
Dh 32 SGNVELENDKRNELINRKEKHPKCSNLSNNTLAPSSVHSPNNSHVNYYSSNNIM 91
Qy 45 DMTKIERHEFTVDKVPDDTVNFTYEN-----TKPLIIDLYENGCVCSFAYGOT 97
Dh 92 DFRKYIK-----DDSSNDITKQSYLSIESHTNM--LSNIYFNFNICN----- 132
Qy 98 GSGKTYMIGSQPYQGSDFRGIFQYAAAGDIFFLNLYDKDNKGFIFISYEIYCGKL-YD 156
Dh 133 -----EINSYFKNP-----LDNTR--VQNNSSSYCQMKYS 163
Qy 157 LQKRKMAVALENGKEVAVVQDKILRYLTKEELIKMIDGLR-----KIGVNSQND 211
Dh 164 DYQHKPCLCHRG-----ILNYSNSENQINSDDVSLKNDYTKISMPFNAS 213
Qy 212 SSRSHAILNIDKDKNKTSLGKI--AFIDLAGSERGADTVSONKQOTGDANINRSL 268
Dh 214 SS-----VNVDSNVVNSNITAPCVNMSSPNMYNKGMSGDIINMQTTMNNLKINDLYL 268
Qy 269 ALKECIRAMDSQKNIIPRDSLETKVLRDIFVGSKSMIANISPTISGCEQTTLTKYS 328
Dh 269 -----NEINNINFDSDI-----SICSHGTS----- 289
Qy 329 SRVKNPKNSCTINBEDPTNTERISILDSKSENNASSIENVYIKSHLNSN--NNNKI 386
Dh 290 ---KSKKHS-----DYSTVPYILSN-----MKPNFTSNVSNNIL 326
Qy 387 NRGKINDKIERN-----ILKVK--SPDKPRGFTST--FGKYSLANDIKIKKKKGL 437
Dh 327 SNMMLPNNIALNMSSSSISINMASCNIPLNNMISTWEOYQILSNLLKAYQOHTSY 386
Qy 438 INYKSTLYNDNTINKKN-----NNNNNNNDNNNDNNNNNNNNSSNV----- 483
Dh 387 VSYDMLIYEHMLINNSNMSNMSNMSNMSNMSNMSNMSNMSNMSNMSNMSNMSN 446
Qy 484 -----NNMI--NHMINN-----INNNIYVNNNNNNNNNNNNNSHNNHP 520
Dh 447 ISNNMSNMLFLKSLISNFIPIHLIPNNSNGKITSNNVNLNFPVSNFTSHMEYNNRK 506
Qy 521 QPNVAFDTSDSFSSLDNN-----CHLANNDSKIF--LH--KKNLRDNIK 563


```

Db 507 KQNLASSKVTLNNNNNSLSSEKLNPLILSSSHIHENOVFNHLPSEHOKNEBEIK 566
Qy 564 NRSCDNIANK--KK-----NNLHL-----AHSVSKLTMTSYDP----- 597
Db 567 NDVYIKHQBQIKTPAYLLNNYYLFPQOKLENQSYTKNMKTDPDPIKENDIINQON 626
Qy 598 -----QKXK-----DNTEFKSNINKMEDNTPPKDILYESRVSNMGNVL--- 636
Db 627 ICNSSMKYNVNNINEKMKCMCHENITEMNISRIKNDKQIV-----VSRSSNVLRQM 681
Qy 637 -----LG- 638
Db 682 IDPSPVNDKIQNESFANTRKAREMNIDICSEAYKSETIKORLSRARKGLEGVGG 741
Qy 639 -----LNKNT-----HHISTKDEHN 655
Db 742 RTEANSQSLSKYDFLNKNTIPNNTNLNRKSKKQNAQJNNSKETCYHIIEDIDSVN 801
Qy 656 D-----NKI-----NNGVINIINNSVNSINNSN--- 679
Db 802 EQITKIDIMNEIKSLKYDENMMRFLCADEKKEBNQNHQEDNIDCANKKVPFNKQSS 861
Qy 680 -----MNSIN-----SNMNSNYSKSNYSNOSISDVQIRYVNMDSNKN--N 722
Db 862 RKRKNNEIKKKHEGKDEKEDNLDKMEKTVEEN---DIEBENITEENITEENYTE 917
Qy 723 DNIFPDASCDNNMTPNITNNR-NNNNNNNNNNNNIDVEN-----YNNRGSTNSMCLY 776
Db 918 ENITEENYTEENITEENITEENITEENITEENITEENITEENITEENITEENITE 977
Qy 777 AYN--SHNFOPDNK--NTSNIOINTNNKQNDGVNVMFCHYNL-----NDKRYL 826
Db 978 LYNEKRIKANKMHNQIKMKSQJNNRNNMNMDEKSSSTDICHSSEMENDERVEVE 1037
Qy 827 ID---LN---NKEQDKNIGHCDN-----NIIORADFEEKKKTNFYNNNNI 867
Db 1038 LDNIKEINGLINKE---NIENDENIKPDKMKKEKAVVITNERKREKREKTCDE--- 1090
Qy 868 VIANNMNMKNS---PMKTYGLCSHTSIDMKNEKNEKMKNEKMKNDHIKSNNSNS 923
Db 1091 ---KNDQCKDNTHEHVEKRYLEEDSEMERNTNNKINVDKNNDEKNNNTDCTCTY 1147
Qy 924 SSSSSNNNY--NNINDDTPFONDY-----CHNDNTPFIRKNNNTINSINYONDI 974
Db 1148 CEDNIKKNITIFLNGVENINDKAKLYSEYECJND---INEKKATIGEOJACTNNDKM 1203
Qy 975 IYTTINSLND-----YMSNTLLHFKEKTYPTLSTNEDIYKMEGKHIRLDDQXY 1025
Db 1204 LYSVAVINDDFMKKVVVEFLSHNDVAIKKELNY-----LNTENKIHREYI-LDNKCTL 1256
Qy 1026 DDNDNNNDV--NNKKNVNDNVNDNNVNDNDKNNVNDN----- 1064
Db 1257 HINDKKNKDIYYNIAKSGIDEMNINSNLFNDYDN--KNNVRDLPRDLVSTDTGMDIIN 1314
Qy 1065 -----NNVNDDDDDVDVFNHKN--FN---NNEYLSYFOKNVDTI----- 1098
Db 1315 KKKKIWLINIGCTINYNDNFVKGSGYSVRYLYNIDLANEYLFKYDKNIDVFLKVLHT 1374
Qy 1099 -----INNCINSLDIS-----MYDTEKILNNIILSKYAKEDNVYKRYINED 1142
Db 1375 CHMKKGLKYLVDKDIASDSCVSYWMDYLLKOLHDEKYEETKETHKQVQK--OD 1431
Qy 1143 IYKMSLEF-IDKTAOSIYKRYLTKLL-----FK---KNVD-TQIN 1182
Db 1432 ENKKNKESINKNTKLEKELVILLESIDLNNKLDNCLCHKKGKVFELAVEIDISRG 1491
Qy 1183 NETSDLRKDLVMCHI--CNNPDDQFHFYAYSRL-----KDIINILIMLROJWCESNIR 1235
Db 1492 NEFKFLRBEGLCHRCN-----VIRPLSTKIGLHKQKQYEL--MHCAITGDR 1539
Qy 1236 LLYOFLV--EYQKNSANSVLANV-----SSNNGDII-LANKKLVQDNK 1277
Db 1540 KLLQNLILYHRKKEYEKKKVKNLKLIHICFGKRYHCFESLPLECTIGCYKCL--KNMK 1597

```

```

Qy 1278 NSMDHNNHKK 1288
Db 1598 MKINNNTIEIK 1608

RESULT 33
T18472
hypothetical protein C0440c - malaria parasite (Plasmodium falciparum)
C/Dates: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C/Accession: T18472
C/Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, November 1998
A/Reference number: Z18937
A/Accession: T18472
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-2269 <LAW>
A/Cross-references: EMBL:AL008970; NID:e1407852; PID:e1332566; PIDN:CAA15615.1
A/Genetics:
A/Map position: 3

Query Match      8.0%; Score 544; DB 2; Length 2269;
Best Local Similarity 21.4%; Pred. No. 2,3e-15;
Matches 273; Conservative 169; Mismatches 327; Indels 504; Gaps 65;

Qy 341 INEEDTNTERISILDSKSEMNASSIENVVYKSNHLSNNNNKINRGKINDKIERNTI 400
Db 1 MNNTFKINKRRRTYVES-----YNIINIANKKRYWC--NDKND-----NNT 42
Qy 401 LKXKSPDKPREGFTSTFGKYSLNDIDKIKN---KKGLIYKSTLYNDNTINKKGN 456
Db 43 INEKTYFNSSSTIKNTANNYK---DNITKSDMCEKRSMDRENIYANNYKNNKRN 98
Qy 457 -NNNNNDNDNDNDNNNNNN--NNDSSVMNMTN--HMNNNNNN-----IN 500
Db 99 YSNKNVYNNIINHDMNRNTYKGDLDGRKQYLSNKHYESKNDENSYNLKKYFNPIPIK 158
Qy 501 VNNNNNNNNNNNSHNNHLPQPNYAFPTDTSDFSLDDMCHLNNNDKSIPLHKKNLRNI 560
Db 159 YNNNNNNNNNNNNNDRONIINNITYKPYD-----DK----- 188
Qy 561 KTKRRSSCDNIMKKNKNLHLARHVSYSKLTWFSYDPQKQNTFFSINKJEDNTPK 620
Db 189 -----NEKSN---KTHSYGNF-----NNEFTNDPKQYIM-SNESN 224
Qy 621 ILYESRVSNMGNVLLGLNKNTHTDISTDENHNDKINNGVNIINNSVNSINNSN- 679
Db 225 MTQGHPIYIHHNNNNNDGRIKN-RKDISL--SSYGVKISNDLSKENTSYIQYNNSSG 281
Qy 680 ---MNSINNSNMNSISYKSNYN--SNOSISDVQIRYVNMEDTSKKNNDNIFPDALSC 732
Db 282 ANRHVIAOHNNNOHNNNOHNNNFNNYHAKAPYQDXY-----KNDSEKFSAYNA 333
Qy 733 -----DNNTYPTNTNN----- 744
Db 334 EOHNDMDKNYQDTHKNNMGHSTKGTALISENGQVFGSFKNKMKERKPFKKKERGSS 393
Qy 745 -----NNNNNNNNNNNNNNIDVENYNNRQDGNNSMCLYVNSHNLPOPD---NNKOTS 793
Db 394 INRGGMNNNNNNNNNNNNY--YNN-----YNNH--PCDAERNNKN--E 434
Qy 794 NIQINNTK--MNQGNVYSNMFCHYLANDK--NYLIDANKEQKQK-----NIHG 841
Db 435 NWCESINSIETEKGDMPI-----YANENCMLKKMIYKDDKKEKIKILITLEDVNN 488
Qy 842 CDNN-----IIONRDEKKKKKTNFYNNNNIVIVNNMGNNS----- 879
Db 489 VQSNITEKLSYKSRNEI-----IMIEKNDVTPIFNGETSGSKSYCEKFLLENIRENK 543
Qy 880 -----PR-----MKYGLCG----- 888

```



```

Db 109 AFDKVFPEADQMLFENSAVPMLEOVL-NGVNCITFAVGQGTGKTYTMSGD-----LSD 163
Qy 116 TPGIFYAAG----DIFPLNIYDKDNTK-GIFISFEIYCGKLYDL---QKRKVAAL 167
Db 164 SDGLISGAGLIPRALYQVLSLDSNOEYAVKCSYIELNIEIRDLVSELRKPARVF 223
Qy 168 EN-----GKKEVVVD---LKILRVLTKEELIKMIDVLRLKIGVNSQNDSS 213
Db 224 EDTSRGQNVVITGIEESYIKVAGDGLRLR-----EGSHRQVAATKCNLDSS 271
Qy 214 RSHAILIIDLK-----DINKYTS---LGKIAFDLASEGADTVSQNKQ 255
Db 272 RSHSIFITLHRYKVSQMTDETNSLTIINNNSDILLRASKLHMDVLASENIGRGAENKR 331
Qy 256 TOTGAMINRSLALKECIRAMDSQKNIHPRDSELTKVLRDIFVGSKSIMINISPTI 315
Db 332 ARETGM-INOSLLTLGRVINALVEKAHHIPYRSKLTIRLDODSLGKTKTSMITVSTN 390
Qy 316 SCCEQTLNLTLYSSRYKVFKNKSTCINEEDPTNTERISI-----LDSKSEMNASSIEN 369
Db 391 TNLLETISTLEVAARAKSIRNKP-----QNNQVFRKVLIKDLVDIRLKNIDLNAATRKKN 446
Qy 370 VV-----IKSNHLSNNNNKINRGIKINDKIERNNI-LKXKSPDKPREGFT 414
Db 447 GVTYLAESTYKELMDRVQNKDLCOEQARKLEVLIDLNVKSSREOLQYVSKSNOEHKKEVEA 506
Qy 415 STEFGYSLLNDIDIKKKKKKGLINYKSTLYNDNTINKKNNNNNNNN-----462
Db 507 LQLOLVNSSTELSEVSKSENEK---LKNELVLEIEKRYKETNNAKTTVAATLSQYRE 562
Qy 463 -----DNNDNNNNNNNNNNSSSVNMMINEM-INNNNNNNIN-----VNNNNN 506
Db 563 SKEVIASLYEKLDRTERNNKENENNFMNLKFNLLTMLRSFHSFTDENGFTILNDPNA 622
Qy 507 NNNNNNSHNNHLPQNYAFTD-TSDPSSLDD-----NCHLANNNDKSFILKKNLRDN 559
Db 623 SMEELLNTHSNQ---LISMTKITEHFOQSIDELQASRSCAVPNS-SIDLIVSELKD- 676
Qy 560 IKLKRRSSCDNIMKKKNNHILARHSVSKLTFMSYDPOKKNKONTFPKSNINRKNEDTPK 619
Db 677 ---SKNSLLDALHSLQDISMSQKLGNGISSELEIQKDMKESY-----R 719
Qy 620 DILYESHNVSNMNGNVLLGLNKNTKTHDISTKDEHNNDKINNGVININNSVNSINNSN 679
Db 720 QLVQELASLYNL-----QHTHEE-----SQKELMYGVARN-----DIDALVKTC 757
Qy 680 MNSINNSNM---NSNSIYKSNYNSNQ---SISDVOIRYVENKMDTSKNNNDNIFDPAIS 731
Db 758 TTSINDADILSDYISDOKSKFESKQODLIANIGKIVSNFLOE-----QNESIYTRA-- 809
Qy 732 CDNNMYFNITNNNNNNNNNNNNNNNNNIDVENYNNRDTGNNSMKLYAVNSHNLPGDNNKN 791
Db 810 -----DILHSLNDTNSNIRKANET-----MNNNSEFFLNA-----841
Qy 792 TSNIQNTINNTKNNODGVNYSMNEFCHYLNLDKNYLIDLNNKEQDKNIHIGCD-----NN 845
Db 842 ASQAEIYKANKERIQKIVENGSQL-----LDSKSKAIHNSNSMSWDHICALALAESQKQVNL 897
Qy 846 IIQNRNDEPKKKKTNFNANN-----IVYVNNMGNNSNPFMKYGLCGSHTSIDNNKNN 899
Db 898 EVQTLDRLLQKVKHSESDNTEKEHQQLDLLESLVGNNDN-----LIDSITPT 945
Qy 900 EMKNNENKDNEMKCNHISKNNNSN--SSSSSSNNNINNNINDDDTPQNDYCHNDNFTIR 957
Db 946 HTLQKTIIDHVKTTSLANTHNLGLGDBESLCNLETTIEDISLVK-----LETTDPT 1000
Qy 958 RKNNTNINSNIYQDDIYITINSLDYMSNTLHFKKXYTPTLSTEDYIYNK 1010
Db 1001 SKRELPAFTPSWTRSSSLIKETTNN--LDSQKFRVRETYTSSNSQNSNPDYDK 1051

```

kinesin-like protein - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
 C/Accession: T38378
 R/McLean, U.; Harris, D.; Barrell, B.G.; Rajadream, M.A.; Walsh, S.V.; Wood, V.
 submitted to the EMBL Data Library, September 1997
 A/Reference number: Z21790
 A/Accession: T38378
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1085 <MCL>
 A/Cross-references: EMBL:Z70691; PIDN:CAA94636.1; GSPDB:GN00066; SPDB:SPAC25G10.07C
 A/Experimental source: clone c25G10
 A/Genetics:
 A/Map position: 1
 C/Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology

Query Match 7.9%; Score 538.5; DB 2; Length 1085;
 Best Local Similarity 20.7%; Pred. No. 1.8e-15;
 Matches 235; Conservative 223; Mismatches 410; Indels 265; Gaps 43;

```

Qy 3 SKTYVVRKRLPSELEKKKSDITTYKNNCTIYID-----EPKYVDMTKYIENHEF 55
Db 71 TNNVVRVVRGRDQD-----VRDNSLAVSTSGAMGALAIQSDPSSMLYTKTY 120
Qy 56 IYVKVFDVTDNTVYENTIKPLIIDLYENGCVCSGCPAGQGTSGKTYTLAGSQPYGQSD 115
Db 121 AFDKVFPEADQMLFENSAVPMLEOVL-NGVNCITFAVGQGTGKTYTMSGD-----LSD 175
Qy 116 TPGIFYAAG----DIFPLNIYDKDNTK-GIFISFEIYCGKLYDL---QKRKVAAL 167
Db 176 SDGLISGAGLIPRALYQVLSLDSNOEYAVKCSYIELNIEIRDLVSELRKPARVF 225
Qy 168 EN-----GKKEVVVD---LKILRVLTKEELIKMIDVLRLKIGVNSQNDSS 213
Db 236 EDTSRGQNVVITGIEESYIKVAGDGLRLR-----EGSHRQVAATKCNLDSS 283
Qy 214 RSHAILIIDLK-----DINKYTS---LGKIAFDLASEGADTVSQNKQ 255
Db 284 RSHSIFITLHRYKVSQMTDETNSLTIINNNSDILLRASKLHMDVLASENIGRGAENKR 343
Qy 256 TOTGAMINRSLALKECIRAMDSQKNIHPRDSELTKVLRDIFVGSKSIMINISPTI 315
Db 344 ARETGM-INOSLLTLGRVINALVEKAHHIPYRSKLTIRLDODSLGKTKTSMITVSTN 402
Qy 316 SCCEQTLNLTLYSSRYKVFKNKSTCINEEDPTNTERISI-----LDSKSEMNASSIEN 369
Db 403 TNLLETISTLEVAARAKSIRNKP-----QNNQVFRKVLIKDLVDIRLKNIDLNAATRKKN 458
Qy 370 VV-----IKSNHLSNNNNKINRGIKINDKIERNNI-LKXKSPDKPREGFT 414
Db 459 GVTYLAESTYKELMDRVQNKDLCOEQARKLEVLIDLNVKSSREOLQYVSKSNOEHKKEVEA 518
Qy 415 STEFGYSLLNDIDIKKKKKKGLINYKSTLYNDNTINKKNNNNNNNN-----462
Db 519 LQLOLVNSSTELSEVSKSENEK---LKNELVLEIEKRYKETNNAKTTVAATLSQYRE 574
Qy 463 -----DNNDNNNNNNNNNDSSSVNMMINEM-INNNNNNNIN-----VNNNNN 506
Db 575 SKEVIASLYEKLDRTERNNKENENNFMNLKFNLLTMLRSFHSFTDENGFTILNDPNA 634
Qy 507 NNNNNNSHNNHLPQNYAFTD-TSDPSSLDD-----NCHLANNNDKSFILKKNLRDN 559
Db 635 SMEELLNTHSNQ---LISMTKITEHFOQSIDELQASRSCAVPNS-SIDLIVSELKD- 688
Qy 560 IKLKRRSSCDNIMKKKNNHILARHSVSKLTFMSYDPOKKNKONTFPKSNINRKNEDTPK 619
Db 689 ---SKNSLLDALHSLQDISMSQKLGNGISSELEIQKDMKESY-----R 731
Qy 620 DILYESHNVSNMNGNVLLGLNKNTKTHDISTKDEHNNDKINNGVININNSVNSINNSN 679
Db 732 QLVQELASLYNL-----QHTHEE-----SQKELMYGVARN-----DIDALVKTC 769

```


C:Genetics:

A:Map position: 3

A:Introns: 56/2; 99/3; 151/3; 310/3; 343/2; 395/3; 419/1; 444/3; 462/3; 489/3; 527/1; 55

A:Note: F1612, 60

C:Superfamily: kinesin-related protein KIP1; kinesin motor domain homology

Query Match

7.8%; Score 534; DB 2; Length 1058;

Best Local Similarity 22.4%; Pred. No. 2.7e-15; Mismatches 412; Indels 184; Gaps 36;

Matches 227; Conservative 192;

```
QY 5 IKVVRKRPISLEKKKSDIITVKNKCTLYIDPRYKDMT-----KYIERHEPIYDK 59
DB 49 VOILIRCPRLSEDEARHPTPVISCNEN-----RREVAATQSIAGKHIDRH-FAPDK 99
QY 60 VEDDTVDNFTYENTIKPLIIDLYENGCVSCFAYGCGSKTYTMTLGSQPYGSDPT-- 117
DB 100 VFGPASQCKDLYDQAIQPIVEVLE-GYNCITFAYGQTGTGTYTMEGARKKNGEFPSPD 158
QY 118 -GIFQYAGDIFTPLINIDKNTKGIFFSFEIYCGKLYDLL-----QKRMVA 165
DB 159 AGVTPRAVKQIFDILEAGAGAYS--MKYTFELEIYNEISDLLAPETIKFYDEKSKSIA 216
QY 166 ALENGKEVVVKDLKILRVLTKEELILKMIDGVLKIGVNSQNDSSSHAILNIDIKD 225
DB 217 LMEDCKGSVFRGLEEEIYSTANEIYKILEKGSARKRTAETLLNKSSRSHSIFSITI-H 275
QY 226 INKNT-----SLGKIAPIDLAGEGADTVSONKQOTOGANINRSLALKECIRAMD 278
DB 276 IKENTPEGEEMIKCGKMLVLDLAGESENISRGAREGRAREAG-ETINKSLTLTGRIYALV 334
QY 279 SDKNHIPEPDSLEIYLRDIFVGSKSIMIANISPTISCCEQTLNLTLYSSRVKNFKNKS 338
DB 335 EHSGHIPYRDSLTFLRESLOGKTKTCVIATISPSHLEETSTLDYAHRAKNIKNKP 394
QY 339 TCINE--EDQNTERRISLDSKSGSEMASSIEEN-VVTKSHLISNNNNKINRGKINDK 394
DB 395 E-INQKMKSAVMKOLYEIDRLKOEYVAAAREKNGIYIPKORYIOEAEKK---AMA EK 449
QY 395 IERNNILKNSKFDKREGFTSPFGKYSLLND--IDIKKKKKGLINYKSTLYNDNTINK 452
DB 450 IERLE-LOSESKDKKVVLDQELYNQOITLAEISEKLEKTEKK-LEETHSLPD--LEB 504
QY 453 KANNNNNNNDNNNDNNNNNNNDSSSMVNNMIMNNNNNNNNNNNNNNNNNNNNNN 512
DB 505 KYRQANATIKK-----EFVISNLKSEKSLVER-----533
QY 513 NSHNNHLPOPNYAFTDSDSSLDPMNCHLNNDKSIPLHKQNLDPNITKLKRRSSCDIM 572
DB 534 -----AFQIRTELESA-----SSDVSNLFSKIERKDKIEDGNRFLOKFO 573
QY 573 NKKKNMLHARHSVSKLT-----MFSYDPOKNKNTFFKSNINMGEDNTPRD 620
DB 574 SOLTOOLELHKTVASSVTQOEVLKHEMEDMESFVSTKSEATEELDRLSKLKR----- 628
QY 621 ILYESRNVSMNNGVNLGLNKVTHHDISTKDENDNDKINNGVINIINN--SNVNSINNS 678
DB 629 -VYGS-GIFALD-NAVAKLDGNSQSTFSSL--NSEVSKSHLELVFKGFASEADMLOD 683
QY 679 NNSNINNSN-----MNSNSIYKSNYSNOSISDVQIRYVNEMDTSKANDNIFPDALS 731
DB 684 LQSSLNKQEEKLITFAQQORAKASRAVDARSVSKTVEFFKTLDTHTATKLTGIVEAQT 743
QY 732 CDNNNYPTNTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 791
DB 744 VNHKKLSEFENKPEECANEBRQLEKVAELLANSNARKKNLVQMAV--HDLRESASTRT 801
QY 792 TSNIQNTNKNKNOG--NVNYSMNFCYVNLNDKYYLIDLNKQKDKNIH---GCDNNI 846
DB 802 TTLQHEMSTMDQSTSIKAEWGI--HMEKTESHHEDTSAYESGKKAQEVILNCKEKT 858
QY 847 IONRNDPEKKKKTFTNNNNIYIVNNMGNNSPRMKYGLCSGHTSIDNMKNEMKNEM 906
DB 859 EMSAHQWRKAQE-----SLVSLERNNV-----ASVDSIVRGMDANE- 895
```

```
QY 907 KDNEMKDNHIXSNNNSSSS-----SNNNIYNNINDDTFONDYCHNDNTEFTI 956
DB 896 -----NIRSQFSTRVSSSLVFPDANSLSLTSIDHSLQLDNDACIKVNSMI 942
```

Search completed: October 2, 2003, 16:21:28
Job time : 94 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: October 2, 2003, 16:17:11 ; Search time 33 Seconds

(without alignments)
1651.406 Million cell updates/sec

Title: US-10-006-780-2

Perfect score: 6842

Sequence: 1 MNSKIKVVRKRLSELEKK.....KKLVQDNKSNMHNHKK 1288

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/ptodata/1/1aa/5A COMB.pep:*

2: /cgn2_6/ptodata/1/1aa/5B COMB.pep:*

3: /cgn2_6/ptodata/1/1aa/6A COMB.pep:*

4: /cgn2_6/ptodata/1/1aa/6B COMB.pep:*

5: /cgn2_6/ptodata/1/1aa/PTUS COMB.pep:*

6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	768.5	11.2	363	4	US-09-967-908A-10
2	768.5	11.2	361	4	US-09-967-908A-6
3	768.5	11.2	1368	4	US-09-967-908A-2
4	758.5	11.1	343	4	US-09-967-908A-8
5	758.5	11.1	373	4	US-09-967-908A-4
6	758.5	11.1	730	4	US-09-281-170A-3
7	758.5	11.1	730	4	US-09-724-884-3
8	752.5	11.0	405	4	US-09-594-669-12
9	752.5	11.0	434	4	US-09-594-669-10
10	752.5	11.0	473	4	US-09-594-669-8
11	752.5	11.0	723	4	US-09-594-669-14
12	750.5	11.0	725	4	US-09-584-669-16
13	750.5	11.0	725	4	US-09-585-684B-33
14	745.5	10.9	362	4	US-09-594-669-6
15	745.5	10.9	331	4	US-09-594-669-4
16	745.5	10.9	430	4	US-09-594-669-2
17	736	10.8	678	4	US-09-595-684B-25
18	696.5	10.2	319	4	US-09-724-510-2
19	696.5	10.2	319	4	US-09-723-216-2
20	696.5	10.2	319	4	US-09-675-227-2
21	589.5	8.6	303	4	US-09-722-129-4
22	589.5	8.6	492	4	US-09-722-129-2
23	561.5	8.2	1111	4	US-09-914-259-28
24	538.5	7.9	1073	4	US-09-541-782-6
25	538.5	7.9	1073	4	US-09-723-820-6
26	538.5	7.9	2184	4	US-09-417-485D-6
27	533.5	7.8	2391	2	US-08-446-855A-2

28	533.5	7.8	2391	3	US-09-150-741-2	Sequence 2, App1
29	500.5	7.3	967	4	US-09-914-259-21	Sequence 21, App1
30	500	7.3	1184	4	US-09-541-782-2	Sequence 2, App1
31	500	7.3	1184	4	US-09-723-820-2	Sequence 2, App1
32	482	7.0	2652	4	US-09-595-684B-31	Sequence 31, App1
33	465.5	6.8	367	4	US-09-724-511-2	Sequence 2, App1
34	465.5	6.8	367	4	US-09-723-097-2	Sequence 2, App1
35	465.5	6.8	367	4	US-09-723-344-2	Sequence 2, App1
36	465	6.8	1038	3	US-09-541-782-4	Sequence 4, App1
37	465	6.8	1038	3	US-09-723-820-4	Sequence 4, App1
38	464.5	6.8	355	4	US-09-724-511-4	Sequence 4, App1
39	464.5	6.8	355	4	US-09-723-097-4	Sequence 4, App1
40	464.5	6.8	355	4	US-09-632-344-4	Sequence 4, App1
41	460.5	6.7	1388	4	US-09-572-191-2	Sequence 2, App1
42	460.5	6.7	1388	4	US-09-723-262-2	Sequence 2, App1
43	460.5	6.7	1388	4	US-09-723-219-2	Sequence 2, App1
44	459	6.7	1066	3	US-09-541-782-8	Sequence 8, App1
45	459	6.7	1066	4	US-09-723-820-8	Sequence 8, App1

ALIGNMENTS

```
RESULT 1
US-09-967-908A-10
; Sequence 10, Application US/09967908A
; Patent No. 6514738
; GENERAL INFORMATION:
; APPLICANT: Cytokine, Inc.
; APPLICANT: Beraud, Christopher
; APPLICANT: Guo, Jun
; APPLICANT: Freedman, Richard
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: KIN-3 MOTOR PROTEIN AND METHODS FOR ITS USE
; FILE REFERENCE: 020552-002400US
; CURRENT APPLICATION NUMBER: US/09/967,908A
; PRIOR APPLICATION NUMBER: 2001-09-28
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide
; OTHER INFORMATION: fragment E213 - S566.
US-09-967-908A-10

Query Match      11.2%; Score 768.5; DB 4; Length 363;
Best Local Similarity 47.5%; Pred. No. 6.4e-49;
Matches 164; Conservative 58; Mismatches 116; Indels 7; Gaps 3;

4 KIKVVRKRLSELEKKKSDIITVKNKCTIYIDPRKVMYTERBHFVDPD 63
12 KIRVCKRKLPLGKREVRREINIIYEDKETLIVHEKKAVDLTQIILQHFVDFEYGE 71
64 TYDNFTYENITKPLIIDYENGCVCSFAYGQSGKTYTM/GSQPYGQSDPTGIFQYA 123
72 ACTNQGVVWKTTHPLQIHFF-NGGNATCFAYGQTGAKTYTMIGTH-----ENPGLYALA 125
124 AGDIFPLFIYDKNTKGIFFISFYEGCKYLDLQKRMVAALENGKEVYVKDKIKR 183
126 ADQIFQLRVSOPRKHLFWISFYELIYCGQLDLNRRRLPARBDSKHMVQIVGIQEQ 185
184 VITKEELIKMDIGVILRKIGVNSQDSSRSHATINDLKQINKTSLGKIAFDPLAS 243
186 VDSVELLEVIILKSGKERSTGATGVNADSSRSHAVIQIYKOSAKRT-FGRISFDPLAS 244
244 EKGADTVSONKQOTQDGANINSLALKECIRAMDSKNHIFPRDSELTQVLADIFVGS 303
```


Db 245 ERADARSDRQRTMEGAELNOSLALKECIRALDQEHHTTTPFROSQTLVQLKDSFGNA 304
QY 304 KSIIMANISPTISCEQTLNLTFRYSRVKPKNSTCINEEDDTN 348
Db 305 KTCMIANISPSHVATEHTLNTLRADRYKELKGIKCTSVTSRN 349

RESULT 2

US-09-967-908A-6
Sequence 6, Application US/09967908A
Patent No. 6514738
GENERAL INFORMATION:
APPLICANT: CytoKinetics, Inc.
APPLICANT: Beraud, Christopher
APPLICANT: Freedman, Richard
APPLICANT: Patel, Umesh A.
APPLICANT: Patel, Umesh A.
TITLE OF INVENTION: KINI-3 MOTOR PROTEIN AND METHODS FOR ITS USE
FILE REFERENCE: 020552-002400US
CURRENT APPLICATION NUMBER: US/09/967,908A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US 09/675,227
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 381
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Peptide
OTHER INFORMATION: fragment V195-5566 with flanking vector sequence.
US-09-967-908A-6

Query Match 11.2%; Score 768.5; DB 4; Length 381;
Best Local Similarity 47.5%; Pred. No. 6.8e-49;
Matches 164; Conservative 58; Mismatches 116; Indels 7; Gaps 3;

QY 4 KIKVVRKRPPLSELEKKKSDITVYKNTLYIDEPRYVDMTKYIERHEFYDKVDD 63
Db 30 KIRVCVRKRPPLGMRERVRGEINIIIVEDKETLLVHEKKEAVDLTOYILQHVFYDEVEFGE 89
QY 64 TVDNFTYENTIKPLIIDLYENGCVCSGFAYGOTSGSKTYTMGSPYGGSDTPGIRQYA 123
Db 90 ACTNODVYMKTHPLIQHIF-NGGNATCFAYGOTGAGKTYTMIGTH-----ENPGLYALA 143
QY 124 AGDIFPLNIYDKONTKGIFFISFEIYCGKLYDLQKRKVAALENGKEVVDLKLIR 183
Db 144 AKDIFRQLEVSQPRKHLFWISFEIYCGQLYDLNRRKRLPAREDSKHWQIVGLQELQ 203
QY 184 VLTRBELILKMDGVLLRKIGVNSQNDSSRSHAILNDLKDINKNTSLGKIAFIDLGS 243
Db 204 VDSVELLEVLILKSKERSGTATGVNADSSRSHAVIQIQKDSAKRT-FGRISFIDLGS 262
QY 244 ERGADTVSONKQOTODGANINRSILALKECIRAMDSQKNHLPDPSELTKVLROIIFYGKS 303
Db 263 ERADARSDRQRTMEGAELNOSLALKECIRALDQEHHTTTPFROSQTLVQLKDSFGNA 322
QY 304 KSIIMANISPTISCEQTLNLTFRYSRVKPKNSTCINEEDDTN 348
Db 323 KTCMIANISPSHVATEHTLNTLRADRYKELKGIKCTSVTSRN 367

RESULT 3

US-09-967-908A-2
Sequence 2, Application US/09967908A
Patent No. 6514738
GENERAL INFORMATION:
APPLICANT: CytoKinetics, Inc.
APPLICANT: Beraud, Christopher
APPLICANT: Guo, Jun
APPLICANT: Freedman, Richard

APPLICANT: Patel, Umesh A.
APPLICANT: Davies, Katherine A.
TITLE OF INVENTION: KINI-3 MOTOR PROTEIN AND METHODS FOR ITS USE
FILE REFERENCE: 020552-002400US
CURRENT APPLICATION NUMBER: US/09/967,908A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US 09/675,227
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1368
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: KINI-3 motor
OTHER INFORMATION: domain fragment
US-09-967-908A-2

Query Match 11.2%; Score 768.5; DB 4; Length 1368;
Best Local Similarity 47.5%; Pred. No. 4.1e-48;
Matches 164; Conservative 58; Mismatches 116; Indels 7; Gaps 3;

QY 4 KIKVVRKRPPLSELEKKKSDITVYKNTLYIDEPRYVDMTKYIERHEFYDKVDD 63
Db 223 KIRVCVRKRPPLGMRERVRGEINIIIVEDKETLLVHEKKEAVDLTOYILQHVFYDEVEFGE 282
QY 64 TVDNFTYENTIKPLIIDLYENGCVCSGFAYGOTSGSKTYTMGSPYGGSDTPGIRQYA 123
Db 283 ACTNODVYMKTHPLIQHIF-NGGNATCFAYGOTGAGKTYTMIGTH-----ENPGLYALA 336
QY 124 AGDIFPLNIYDKONTKGIFFISFEIYCGKLYDLQKRKVAALENGKEVVDLKLIR 183
Db 337 AKDIFRQLEVSQPRKHLFWISFEIYCGQLYDLNRRKRLPAREDSKHWQIVGLQELQ 396
QY 184 VLTRBELILKMDGVLLRKIGVNSQNDSSRSHAILNDLKDINKNTSLGKIAFIDLGS 243
Db 397 VDSVELLEVLILKSKERSGTATGVNADSSRSHAVIQIQKDSAKRT-FGRISFIDLGS 455
QY 244 ERGADTVSONKQOTODGANINRSILALKECIRAMDSQKNHLPDPSELTKVLROIIFYGKS 303
Db 456 ERADARSDRQRTMEGAELNOSLALKECIRALDQEHHTTTPFROSQTLVQLKDSFGNA 515
QY 304 KSIIMANISPTISCEQTLNLTFRYSRVKPKNSTCINEEDDTN 348
Db 516 KTCMIANISPSHVATEHTLNTLRADRYKELKGIKCTSVTSRN 560

RESULT 4

US-09-967-908A-8
Sequence 8, Application US/09967908A
Patent No. 6514738
GENERAL INFORMATION:
APPLICANT: CytoKinetics, Inc.
APPLICANT: Beraud, Christopher
APPLICANT: Guo, Jun
APPLICANT: Freedman, Richard
APPLICANT: Patel, Umesh A.
APPLICANT: Davies, Katherine A.
TITLE OF INVENTION: KINI-3 MOTOR PROTEIN AND METHODS FOR ITS USE
FILE REFERENCE: 020552-002400US
CURRENT APPLICATION NUMBER: US/09/967,908A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US 09/675,227
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 343
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Peptide

OTHER INFORMATION: fragment E213-S546
US-09-967-908A-8

Query Match	11.1%;	Score 758.5;	DB 4;	Length 343;
Best Local Similarity	48.9%;	Pred. No. 3.2e-48;		
Matches 161;	Conservative	58;	Mismatches 103;	Indels 7;
				Gaps 3

[illegible][illegible]

RESULT 6
US-09-291-170A-3
Sequence 3, Application US/09291170A
Patent No. 6410687
GENERAL INFORMATION:
APPLICANT: Vale, Ronald D.
APPLICANT: Hartman, James J.
TITLE OF INVENTION: Assays for the University of California
TITLE OF INVENTION: Depolymerization Inhibitors
FILE REFERENCE: 185578-000510US
CURRENT APPLICATION NUMBER: US/09/291,170A
CURRENT FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 60/081,734
PRIOR FILING DATE: 1998-04-14
NUMBER OF SEQ. ID NOS: 16
SOFTWARE: PatentIn Ver. 2.0
SEQ. ID NO 3
LENGTH: 730
TYPE: PR1
ORGANISM: *Xenopus laevis*
FEATURE:
OTHER INFORMATION: *Xenopus* kinesin central motor 1 (XKCM1)
US-09-291-170A-3

[illegible]

RESULT 7
US-09-724-884-3
; Sequence 3, Application US/09724884
; Patent No. 6429304
; GENERAL INFORMATION:
; APPLICANT: Vale, Ronald D.
; APPLICANT: Hartman, James J.
; TITLE OF INVENTION: Assays for the Detection of Microtubule
; FILE OF INVENTION: Depolymerization Inhibitors
; FILE REFERENCE: 18557B-000510US
; CURRENT APPLICATION NUMBER: US/09/724,884
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/291,170
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Xenopus laevis
; FEATURE:
; OTHER INFORMATION: Xenopus kinesin central motor 1 (XKCM1)
US-09-724-884-3

Query Match 11.1%; Score 758.5; DB 4; Length 730;
Best Local Similarity 42.1%; Pred. No. 9.3e-48;
Matches 175; Conservative 84; Mismatches 130; Indels 27; Gaps 12;

QY 4 KIKVVRKRPRLSELEKKKSDIITVKNKCTLYIDEPYKVDMTKYEIRHEFIYDKVFD 63
Db 262 RICVCKRPLKQELAKKEIDIVISPSKLVVHEPKLTKYLENQAFCDFAFDE 321

QY 64 TVDNFTYVENTIKPLIIDLYENGCVCSFAYGOTGSGKTYTMLS-QPYGSDTPGIFQY 122
Db 322 TATNEVYRFTARPLVQTIIEGG-KATCFAYGOTGSGKTYTMLS-QPYGSDTPGIFQY 380

QY 123 AAGDIFFTFLN-IYDKONTGIFISPEYICGLYLLQKRVNVALENGKKEVYVDLX 178
Db 381 ASRDVFLKNQPCYRKLGLG-VYTFPEIYNGKLFIDLKAKLRVLEDKQOVVGLQ 437

QY 179 LKIRLVLTKEELILKID-GVLLRKIGVNSONDESSRSHAIINTDLKDKINKTSL-GKIA 236
Db 438 LLEKQVLSAD-VERKMEIGSACRTSGOTFANTSSRSRSHACQIILR-RGSKLHGKFS 493

QY 237 FIDLAGSERGADTVSONKQOTGDGANINRSLLALKECIRAMSDKNHIFPDSBELTVLR 296
Db 494 LVDLAGNERGVDTASADRITMKGAENRSLLALKECIRALGQNSHTPPRESKLTQILR 553

QY 297 DIFPGK-SKSMINANISPTISCCBOTLNTIRYSRYKFNKSTCINEEDTTERISIL 355
Db 554 DSFGNSRMTMIMLSPGFNSCEYTLNTLRVADRVKLSPOA-----BTNDNLOME 607

QY 356 DSKSEMAVASSIENVVKSNNHLLSNNNKINRGKINDKIERNNILKNSFDPKPRE 411
Db 608 DSGSH--ASIEGLQLODFLKDDELSHN--SFQALNRVGLERKAVDELRE 658

RESULT 8
US-09-594-669-12
; Sequence 12, Application US/09594669
; Patent No. 6331424
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: No. 6331424e1 motor proteins and methods for
; FILE OF INVENTION: their use
; FILE REFERENCE: 1042
; CURRENT APPLICATION NUMBER: US/09/594,669
; CURRENT FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/295,612

; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: US 09/314,464
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Human
US-09-594-669-12

Query Match 11.0%; Score 752.5; DB 4; Length 405;
Best Local Similarity 42.1%; Pred. No. 1.1e-47;
Matches 174; Conservative 75; Mismatches 143; Indels 21; Gaps 10;

QY 4 KIKVVRKRPRLSELEKKKSDIITVKNKCTLYIDEPYKVDMTKYEIRHEFIYDKVFD 63
Db 3 RICVCKRPLKQELAKKEIDIVISPSKLVVHEPKLTKYLENQAFCDFAFDE 62

QY 64 TVDNFTYVENTIKPLIIDLYENGCVCSFAYGOTGSGKTYTMLS-QPYGSDTPGIFQY 122
Db 63 TATNEVYRFTARPLVQTIIEGG-KATCFAYGOTGSGKTYTMLS-QPYGSDTPGIFQY 121

QY 123 AAGDIFFTFLN-IYDKONTGIFISPEYICGLYLLQKRVNVALENGKKEVYVDLX 180
Db 122 ASRDVFLKNQPCYRKLGLG-VYTFPEIYNGKLFIDLKAKLRVLEDKQOVVGLQ 180

QY 181 ILVLVTKEELILKID-GVLLRKIGVNSONDESSRSHAIINTDLKDKINKTSL-GKIAFID 239
Db 181 EHLVNSAD-VERKMEIGSACRTSGOTFANTSSRSRSHACQIILR-AKGRMHKESLVD 237

QY 240 LAGSERGADTVSONKQOTGDGANINRSLLALKECIRAMSDKNHIFPDSBELTVLRDIF 299
Db 238 LAGNERGADTVSONKQOTGDGANINRSLLALKECIRALGQNSHTPPRESKLTQILRDSF 297

QY 300 VGR-SKSMINANISPTISCCBOTLNTIRYSRYKFNKSTCINEEDTTERISILDSK 358
Db 298 IGENSRMTMIMLSPGFNSCEYTLNTLRVADRVKLSPOA-----BTNDNLOME 348

QY 359 GSEMAVASSIENVVKSNNHLLSNNNKINRGKINDKIERNNILKNSFDPKPRE 411
Db 349 TEEMKES--NGALIPNLSEBELSSONSFEANATQIREKAMEELKE 399

RESULT 9
US-09-594-669-10
; Sequence 10, Application US/09594669
; Patent No. 6331424
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: No. 6331424e1 motor proteins and methods for
; FILE OF INVENTION: their use
; FILE REFERENCE: 1042
; CURRENT APPLICATION NUMBER: US/09/594,669
; CURRENT FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: US 09/314,464
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Human
US-09-594-669-10

Query Match 11.0%; Score 752.5; DB 4; Length 434;
Best Local Similarity 42.1%; Pred. No. 1.2e-47;
Matches 174; Conservative 75; Mismatches 143; Indels 21; Gaps 10;

QY 4 KIKVVRKRPRLSELEKKKSDIITVKNKCTLYIDEPYKVDMTKYEIRHEFIYDKVFD 63


```

Db      32 RLCVCRKRLPKOELAKKEIDIVISIPSKCLLVHPEKLVLDLTKYLENCAFCEFAFDE 91
Qy      64 TYDNFTYVENTIKPLIDLYENGCVSCFAYGOTSGKTYTMLS-QPGOSDTPGIFQY 122
Db      92 TASNEVYRFTARPLVOTIFEGG-KATCFAYGOTSGKTHMGDLSGKAQNSKGIYAM 150
Qy      123 AAGDIFTFLN--IYDKONTKGIFISFEYIYCGKLYDLQKRKVAALENGKEVVVDLK 180
Db      151 ASRDVFLKQPCYRKLGLE-VYTFEYIYNGKLFIDLNKAKRLVLEDEKQOVVGLQ 209
Qy      181 ILKVLTKBELIKMID-GVLLRKIGVNSQNDSSRSRSHALINDLKQINKNTSLGKIAFID 239
Db      210 EHLVNSADD-VIKMIDWGSACRTSGQTFANSNSRSHACFOILLR--AKRWHGKFSYLD 266
Qy      240 LAGSERGADTVSONKQOTODGANINRSLALKECIRAMSDKNHIFRSELTQVLARDIF 299
Db      267 LAGNERGADTVSSADQOTRMGAELNLSLALKECIRALGONKATPFRSKLTQVLARDISF 326
Qy      300 VGR-SKSIMIANISPTISCEQOTLNTLRYSRVKFNKSTCINEEDDTNTERTISLDSK 358
Db      327 IGNSRRTCMATISPGISSCEYTLNTLRVADRVKELSPHSGPGEQ-----LIQME 377
Qy      359 GSEMASSIEVNVVIKSNHLSNNNNKINRGKINDKIERNNILKNKSPDKPRE 411
Db      378 TEEMERACS--NGALIPGNLSKEEBELSSQMSFNEAMTQIRELEKAMEELKE 428

```

RESULT 10

```

US-09-594-669-8
; Sequence 8, Application US/09594669
; Patent No. 6331424
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: No. 6331424e1 motor proteins and methods for
; FILE REFERENCE: 1042
; CURRENT APPLICATION NUMBER: US/09/594,669
; CURRENT FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: US 09/314,464
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Human
US-09-594-669-8

```

```

Query Match      11.0%; Score 752.5; DB 4; Length 473;
Best Local Similarity 42.1%; Pred. No. 1,4e-47;
Matches 174; Conservative 75; Mismatches 143; Indels 21; Gaps 10;

```

```

Qy      4 KIKVVVRKRPPLSELEKKKDSOIIYKNNCTLYIDPRYKVDMTKIERHEFIYDKVFPD 63
Db      71 RLCVCRKRLPKOELAKKEIDIVISIPSKCLLVHPEKLVLDLTKYLENCAFCEFAFDE 130
Qy      64 TYDNFTYVENTIKPLIDLYENGCVSCFAYGOTSGKTYTMLS-QPGOSDTPGIFQY 122
Db      131 TASNEVYRFTARPLVOTIFEGG-KATCFAYGOTSGKTHMGDLSGKAQNSKGIYAM 189
Qy      123 AAGDIFTFLN--IYDKONTKGIFISFEYIYCGKLYDLQKRKVAALENGKEVVVDLK 180
Db      190 ASRDVFLKQPCYRKLGLE-VYTFEYIYNGKLFIDLNKAKRLVLEDEKQOVVGLQ 248
Qy      181 ILKVLTKBELIKMID-GVLLRKIGVNSQNDSSRSRSHALINDLKQINKNTSLGKIAFID 239
Db      249 EHLVNSADD-VIKMIDWGSACRTSGQTFANSNSRSHACFOILLR--AKRWHGKFSYLD 305
Qy      240 LAGSERGADTVSONKQOTODGANINRSLALKECIRAMSDKNHIFRSELTQVLARDIF 299

```

```

Db      306 LAGNERGADTVSSADQOTRMGAELNLSLALKECIRALGONKATPFRSKLTQVLARDISF 365
Qy      300 VGR-SKSIMIANISPTISCEQOTLNTLRYSRVKFNKSTCINEEDDTNTERTISLDSK 358
Db      366 IGNSRRTCMATISPGISSCEYTLNTLRVADRVKELSPHSGPGEQ-----LIQME 416
Qy      359 GSEMASSIEVNVVIKSNHLSNNNNKINRGKINDKIERNNILKNKSPDKPRE 411
Db      417 TEEMERACS--NGALIPGNLSKEEBELSSQMSFNEAMTQIRELEKAMEELKE 467

```

RESULT 11

```

US-09-594-669-14
; Sequence 14, Application US/09594669
; Patent No. 6331424
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: No. 6331424e1 motor proteins and methods for
; FILE REFERENCE: 1042
; CURRENT APPLICATION NUMBER: US/09/594,669
; CURRENT FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: US 09/314,464
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Human
US-09-594-669-14

```

```

Query Match      11.0%; Score 752.5; DB 4; Length 723;
Best Local Similarity 42.1%; Pred. No. 2,6e-47;
Matches 174; Conservative 75; Mismatches 143; Indels 21; Gaps 10;

```

```

Qy      4 KIKVVVRKRPPLSELEKKKDSOIIYKNNCTLYIDPRYKVDMTKIERHEFIYDKVFPD 63
Db      256 RLCVCRKRLPKOELAKKEIDIVISIPSKCLLVHPEKLVLDLTKYLENCAFCEFAFDE 315
Qy      64 TYDNFTYVENTIKPLIDLYENGCVSCFAYGOTSGKTYTMLS-QPGOSDTPGIFQY 122
Db      316 TASNEVYRFTARPLVOTIFEGG-KATCFAYGOTSGKTHMGDLSGKAQNSKGIYAM 374
Qy      123 AAGDIFTFLN--IYDKONTKGIFISFEYIYCGKLYDLQKRKVAALENGKEVVVDLK 180
Db      375 ASRDVFLKQPCYRKLGLE-VYTFEYIYNGKLFIDLNKAKRLVLEDEKQOVVGLQ 433
Qy      181 ILKVLTKBELIKMID-GVLLRKIGVNSQNDSSRSRSHALINDLKQINKNTSLGKIAFID 239
Db      434 EHLVNSADD-VIKMIDWGSACRTSGQTFANSNSRSHACFOILLR--AKRWHGKFSYLD 490
Qy      240 LAGSERGADTVSONKQOTODGANINRSLALKECIRAMSDKNHIFRSELTQVLARDIF 299
Db      491 LAGNERGADTVSSADQOTRMGAELNLSLALKECIRALGONKATPFRSKLTQVLARDISF 550
Qy      300 VGR-SKSIMIANISPTISCEQOTLNTLRYSRVKFNKSTCINEEDDTNTERTISLDSK 358
Db      551 IGNSRRTCMATISPGISSCEYTLNTLRVADRVKELSPHSGPGEQ-----LIQME 601
Qy      359 GSEMASSIEVNVVIKSNHLSNNNNKINRGKINDKIERNNILKNKSPDKPRE 411
Db      602 TEEMERACS--NGALIPGNLSKEEBELSSQMSFNEAMTQIRELEKAMEELKE 652

```

RESULT 12

```

US-09-594-669-16
; Sequence 16, Application US/09594669
; Patent No. 6331424

```



```
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: No. 6331424el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1042
; CURRENT APPLICATION NUMBER: US/09/594,669
; CURRENT FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: US 09/314,464
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 725
; TYPE: PRT
; ORGANISM: Human
US-09-594-669-16

Query Match      11.0%; Score 750.5; DB 4; Length 725;
Best Local Similarity 41.9%; Pred. No. 3.6e-47;
Matches 173; Conservative 76; Mismatches 143; Indels 21; Gaps 10;

Qy      4 KIKVVRKRPPLSELEKKKKSDIITVKNKCTLYIDEPYKVDMTKYIERHFIYDKVPDD 63
Db      258 RICCVKRPPLNKDELAKKEIDVISIPSKCLLVHBPKLKYLDKYLENQAFCDFPADE 317
Qy      64 TVNFTYVENTIKPLIIDLYENGCVSCFAYGQTSGKTYMTLGS-QPYGSDTPGIFQY 122
Db      318 TASNEVVYRFTAPLVOTTFEGG-KATCFAYGQTSGKTHMGDLSGKQANASKGIYAM 376
Qy      123 AAGDIPTFLN--IYDKNTKGFISFYIYCGKLYDLQKKNVAALENGKEVYVNDLK 180
Db      377 ASRDVFLKQPCYRKLGLE-VYTFEYIYNGKLPDLNKKAKRLVLEDKQOVVGLQ 435
Qy      181 ILRVLTKEELIKKID-GVLLRKIGVNSQNDSSRSRSHAILNIDLKINKTSLGKIAFID 239
Db      436 EHLVNSADD-VIKMLDWSACRTSGQTFANSSRSRSHACFOIILR--AKGRMHGKESLVD 492
Qy      240 LAGSERGADTVSONKOTQTDGANINRSLALKECIRAMDSKNHIFPDSSELTIVLRDIF 299
Db      493 LAGNERGADTSSADROTMEGAELIKSLALKECIRALGQNKATPFRESKLTQVLRDSF 552
Qy      300 VGR-SKSMIANISPTISCCQOTLNTLYSRVKNFKKSTCINEDDTNTERISIIDSX 358
Db      553 IGENSRMTMIATISPGISSCEYTLNLTLYADRVKELSPHSGPSGEO-----LIQME 603
Qy      359 GSEMASSIEENVVIKSNHLNNSNNKINRGINDKIERNNILKKSFDKPRE 411
Db      604 TEEMEACS--NGALIPGNLSKEEELSSQMSFNEAMTOIRELEKAMEBLKE 654

RESULT 13
US-09-595-684B-33
; Sequence 33, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Valsberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; TITLE OF INVENTION: and purifying human kinesins
; FILE REFERENCE: cytop036
; CURRENT APPLICATION NUMBER: US/09/595,664B
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 33
; LENGTH: 725
; TYPE: PRT
; ORGANISM: Human
US-09-595-684B-33

Query Match      11.0%; Score 750.5; DB 4; Length 725;
Best Local Similarity 41.9%; Pred. No. 3.6e-47;
Matches 173; Conservative 76; Mismatches 143; Indels 21; Gaps 10;

Qy      4 KIKVVRKRPPLSELEKKKKSDIITVKNKCTLYIDEPYKVDMTKYIERHFIYDKVPDD 63
Db      258 RICCVKRPPLNKDELAKKEIDVISIPSKCLLVHBPKLKYLDKYLENQAFCDFPADE 317
Qy      64 TVNFTYVENTIKPLIIDLYENGCVSCFAYGQTSGKTYMTLGS-QPYGSDTPGIFQY 122
Db      318 TASNEVVYRFTAPLVOTTFEGG-KATCFAYGQTSGKTHMGDLSGKQANASKGIYAM 376
Qy      123 AAGDIPTFLN--IYDKNTKGFISFYIYCGKLYDLQKKNVAALENGKEVYVNDLK 180
Db      377 ASRDVFLKQPCYRKLGLE-VYTFEYIYNGKLPDLNKKAKRLVLEDKQOVVGLQ 435
Qy      181 ILRVLTKEELIKKID-GVLLRKIGVNSQNDSSRSRSHAILNIDLKINKTSLGKIAFID 239
Db      436 EHLVNSADD-VIKMLDWSACRTSGQTFANSSRSRSHACFOIILR--AKGRMHGKESLVD 492
Qy      240 LAGSERGADTVSONKOTQTDGANINRSLALKECIRAMDSKNHIFPDSSELTIVLRDIF 299
Db      493 LAGNERGADTSSADROTMEGAELIKSLALKECIRALGQNKATPFRESKLTQVLRDSF 552
Qy      300 VGR-SKSMIANISPTISCCQOTLNTLYSRVKNFKKSTCINEDDTNTERISIIDSX 358
Db      553 IGENSRMTMIATISPGISSCEYTLNLTLYADRVKELSPHSGPSGEO-----LIQME 603
Qy      359 GSEMASSIEENVVIKSNHLNNSNNKINRGINDKIERNNILKKSFDKPRE 411
Db      604 TEEMEACS--NGALIPGNLSKEEELSSQMSFNEAMTOIRELEKAMEBLKE 654

RESULT 14
US-09-594-669-6
; Sequence 6, Application US/09594669
; Patent No. 6331424
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: No. 6331424el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1042
; CURRENT APPLICATION NUMBER: US/09/594,669
; CURRENT FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: US 09/314,464
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Human
US-09-594-669-6

Query Match      10.9%; Score 745.5; DB 4; Length 362;
Best Local Similarity 45.7%; Pred. No. 3.2e-47;
Matches 168; Conservative 65; Mismatches 116; Indels 19; Gaps 9;

Qy      4 KIKVVRKRPPLSELEKKKKSDIITVKNKCTLYIDEPYKVDMTKYIERHFIYDKVPDD 63
Db      3 RICCVKRPPLNKDELAKKEIDVISIPSKCLLVHBPKLKYLDKYLENQAFCDFPADE 62
Qy      64 TVNFTYVENTIKPLIIDLYENGCVSCFAYGQTSGKTYMTLGS-QPYGSDTPGIFQY 122
```



```

/ APPLICANT: Wood, Kenneth
/ APPLICANT: Yu, Ming
/ TITLE OF INVENTION: human kinesins and methods of producing
/ TITLE OF INVENTION: and purifying human kinesins
/ FILE REFERENCE: cytop036
/ CURRENT APPLICATION NUMBER: US/09/595,684B
/ PRIOR FILING DATE: 2002-06-24
/ PRIOR APPLICATION NUMBER: 09/293,612
/ PRIOR FILING DATE: 2000-04-20
/ NUMBER OF SEQ ID NOS: 105
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 25
/ LENGTH: 678
/ TYPE: PRT
/ ORGANISM: Human
US-09-595-684B-25

Query Match 10.8% Score 736; DB 4; Length 678;
Best Local Similarity 45.6%; Pred.No.3.9e-46;
Matches 166; Conservative 62; Mismatches 112; Indels 24; Gaps 8

QY 4 KIKVVRRLPSELEKKKKDDIITVKNNCTLYIDEPKYVDMTKYIRRHEFIYDKVPDD 63
DB 196 RICVCRRRLPKKETQKDDIVITPSKDVMMHPRKQKDLTYLLENQFRPFYAPDD 255
QY 64 TVDNFTYENTIKPLIIDLYENGVCSCFAYGQTSKGITYMLGS-QPYGSDTPGIFQY 122
DB 256 SAPNEMVYRFPAKPLVETIFERG-WATCFAYGQTSKGTHMGDPFGSKNDQCSKGIYAL 314
QY 123 AAGDIFTPLNIYDKONTK----GFIISFYEIYCGYLXLLQKRNVALLENCKEYVVKD 178
DB 315 AARVFLMLK--KENVKLEIQVATFEFISGVFDLNRKTKLRYLEDGKQOVAVG 371
QY 179 LKIRLVLTKEELIKMIDGVLLRKIGVNSQNDSSRSRAIINIDLDKINKNTSLGIAFI 238
DB 372 LQERVKCVED-VLKIIDINSRTSGQISANAHSSRSRAVPIILR--RKGLHKKPSLI 428
QY 239 DLAGSERGADTVSONKQOTQTDGANINRSILALKECIRAMDSDKNIIPRDESELTQVLDI 298
DB 429 DLAGNERGADTVSSADROTRLGAEIINKSLALKECIRALGRNKPPTPPRSKLTQVLRDS 488
QY 489 FVGG-SKSIIMANISFPTISCCGQTLNTRYSRRVKNPKKSKTCT-----NEEDD 346
DB 347 TQTE 350
QY 549 LETQ 552
DB 549 LETQ 552

RESULT 18
US-09-724-510-2
/ Sequence 2, Application US/09724510
/ Patent No. 6432659
/ GENERAL INFORMATION:
/ APPLICANT: Berard, Christophe
/ APPLICANT: Freedman, Richard
/ TITLE OF INVENTION: No. 6432659e1 motor proteins and methods for
/ TITLE OF INVENTION: their use
/ FILE REFERENCE: 1032
/ CURRENT APPLICATION NUMBER: US/09/724,510
/ CURRENT FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 09/675,227
/ PRIOR FILING DATE: 2000-09-29
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 319
/ TYPE: PRT
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: VARIANT
/ LOCATION: (154)...(183)

```

```

: OTHER INFORMATION: Xaa = any amino acid
US-09-724-510-2

Query Match      10.2%; Score 696.5; DB 4; Length 319;
Best Local Similarity 45.8%; Pred. No. 1.1e-43;
Matches 149; Conservative 51; Mismatches 118; Indels 7; Gaps 3

QY      4 KIKVVRKRPSELEKKKKDDITTVKNNCTLYIDEPKVKDMTYIERHEFIVDKVPDD 63
      1 KIRCVRRKRPFGMEVRGRGEINIIIVKEKTLIVHEKKEAVDLTOYIIQHVFYFDEVRGE 60
DB      1 TVDNFTYVENTIKPPIIDLYENGCVCSCFAYAGQTSGGTTYTMLGSOPYGQSDPTPGIFOYA 123
      61 ACTNODVYMKTKTHPIQHIF--NGGNATCFAGYQAGKTYTMIGTH-----ENPGIYALA 114
QY      124 AGDIFFTPLNIYDKNTKGIFFISFYEYIGKLYDLIQKRMAALENGKEVVVKDLKILR 183
      115 AKDIFRQLEVSQPRKHLFWWISFEYIGQLYDLILNRRKXXXXXXXKXXXXXXXXXXXXX 174
DB      184 VLTEELLKIMIDGVLLKRIKIVNSQNDSSSHALINLDLDINQTSIGKIAFIIDLAGS 243
      175 XXXXXXXXXVILKSKESKERTAGTGVNADSSSHAVIQIQKDSART--FGRISFDLAGS 233
QY      244 ERGADTVSONKOTQDGNINRSILATKECTRAMDSKXHIIPRDSLETKVLRIFFVGS 303
      234 ERADARDSDQTKMEGAEINOSLALKECTRALDQETHTFPRQSKLTQVLKDSFIGNA 293
DB      304 KSIMANISPTISCCQTLNLTLYRS 328
      294 KTCMIANISPSHVATHTLTLRYA 318

RESULT 19
US-09-723-216-2
: Sequence 2, Application US/09723216
: Patent No. 6436686
: GENERAL INFORMATION:
: APPLICANT: Beraud, Christophe
: TITLE OF INVENTION: No. 6436686el motor proteins and methods for
: TITLE OF INVENTION: their use
: FILE REFERENCE: 1032
: CURRENT APPLICATION NUMBER: US/09/723,216
: CURRENT FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 09/675,227
: PRIOR FILING DATE: 2000-09-29
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 319
: TYPE: PRT
: ORGANISM: Human
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: (154)...(183)
: OTHER INFORMATION: Xaa = any amino acid
US-09-723-216-2

Query Match      10.2%; Score 696.5; DB 4; Length 319;
Best Local Similarity 45.8%; Pred. No. 1.1e-43;
Matches 149; Conservative 51; Mismatches 118; Indels 7; Gaps 3

QY      4 KIKVVRKRPSELEKKKKDDITTVKNNCTLYIDEPKVKDMTYIERHEFIVDKVPDD 63
      1 KIRCVRRKRPFGMEVRGRGEINIIIVKEKTLIVHEKKEAVDLTOYIIQHVFYFDEVRGE 60
DB      1 TVDNFTYVENTIKPPIIDLYENGCVCSCFAYAGQTSGGTTYTMLGSOPYGQSDPTPGIFOYA 123
      61 ACTNODVYMKTKTHPIQHIF--NGGNATCFAYQAGKTYTMIGTH-----ENPGIYALA 114
QY      124 AGDIFFTPLNIYDKNTKGIFFISFYEYIGKLYDLIQKRMAALENGKEVVVKDLKILR 183
      115 AKDIFRQLEVSQPRKHLFWWISFEYIGQLYDLILNRRKXXXXXXXXXXXXXKXXXXXXXXXXXXX 174
DB      184 VLTEELLKIMIDGVLLKRIKIVNSQNDSSSHALINLDLDINQTSIGKIAFIIDLAGS 243
      175 XXXXXXXXXVILKSKESKERTAGTGVNADSSSHAVIQIQKDSART--FGRISFDLAGS 233
QY      244 ERGADTVSONKOTQDGNINRSILATKECTRAMDSKXHIIPRDSLETKVLRIFFVGS 303
      234 ERADARDSDQTKMEGAEINOSLALKECTRALDQETHTFPRQSKLTQVLKDSFIGNA 293
DB      304 KSIMANISPTISCCQTLNLTLYRS 328
      294 KTCMIANISPSHVATHTLTLRYA 318

```



```

QY      184 VLTKEELIKMDGVLLRKIGVNSONDESSRSHAIINIDKIDINKNTSLGKIAFIDLAGS 243
      175 XXXXXXXXXXXXVILNGSKERSGTGAGVADSSRSHAVIQIQKDSAKRT-FGRISFIDLAGS 233
QY      244 ERGADTVSONKQOTQTDGANINRSLALKECIRAMSDSKNHIIPRDSLTLYLADIVGKS 303
      234 ERADADARSDRQRTMGAEINOSILALKECIRALDQHTHTPPROSKLTQVLKDSFGIGNA 293
QY      304 KSIMINISPTISCCBOTLNTLRYS 328
      294 KTCMIANISPSHVATHTLNTLRKYA 318

```

```

RESULT 20
US-09-675-227-2
; Sequence 2, Application US/09675227
; Patent No. 6461855
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6461855el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1032
; CURRENT APPLICATION NUMBER: US/09/675,227
; CURRENT FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (154)...(183)
; OTHER INFORMATION: Xaa = any amino acid
US-09-675-227-2

```

```

Query Match      10.2%; Score 696.5; DB 4; Length 319;
Best Local Similarity 45.8%; Pred. No. 1.1e-43;
Matches 149; Conservative 51; Mismatches 118; Indels 7; Gaps 3;

```

```

QY      4 KIKVVRKRPPLSELEKKKDSIIITVKNCTLYIDPRYKVDMTKYIERHEFIVDKVPDD 63
      1 KIRVCVRKRPPLNQREBETTLKDLITVPSDNNVWVHESKQVLDLTRYLQNTQFCFHAFFD 63
DB      1 KIRVCVRKRPPLNQREBETTLKDLITVPSDNNVWVHESKQVLDLTRYLQNTQFCFHAFFD 60
QY      64 TVDNFTVYENTIKPLIIDLYENGCVSCFAYGQTSCKTYTMLGSOPYGSDTPGIFQYA 123
      61 ACTNODVYMKTTAPLIQHIF-NGGNATCFAYGQTSCKTYTMTGTH----ENPGLYALA 114
DB      61 ACTNODVYMKTTAPLIQHIF-NGGNATCFAYGQTSCKTYTMTGTH----ENPGLYALA 114
QY      124 AGDIFTEFLNIDYDKNTKGFISFEYIYCGKLYDLOKRKMVALENGKKEVVVVKDKILR 183
      115 AKDIFPOLEYSQPKRLFWISFEYIYCGQYDLNRRKXXXXXXXXXXXXXXXXXXXX 174
DB      115 AKDIFPOLEYSQPKRLFWISFEYIYCGQYDLNRRKXXXXXXXXXXXXXXXXXXXX 174
QY      184 VLTKEELIKMDGVLLRKIGVNSONDESSRSHAIINIDKIDINKNTSLGKIAFIDLAGS 243
      175 XXXXXXXXXXXXVILNGSKERSGTGAGVADSSRSHAVIQIQKDSAKRT-FGRISFIDLAGS 233
DB      175 XXXXXXXXXXXXVILNGSKERSGTGAGVADSSRSHAVIQIQKDSAKRT-FGRISFIDLAGS 233
QY      244 ERGADTVSONKQOTQTDGANINRSLALKECIRAMSDSKNHIIPRDSLTLYLADIVGKS 303
      234 ERADADARSDRQRTMGAEINOSILALKECIRALDQHTHTPPROSKLTQVLKDSFGIGNA 293
DB      234 ERADADARSDRQRTMGAEINOSILALKECIRALDQHTHTPPROSKLTQVLKDSFGIGNA 293
QY      304 KSIMINISPTISCCBOTLNTLRYS 328
      294 KTCMIANISPSHVATHTLNTLRKYA 318

```

```

RESULT 21
US-09-722-129-4
; Sequence 4, Application US/09722129
; Patent No. 6582958
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe

```

```

; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6582958el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1054
; CURRENT APPLICATION NUMBER: US/09/722,129
; CURRENT FILING DATE: 2000-11-24
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Human
US-09-722-129-4

```

```

Query Match      8.6%; Score 589.5; DB 4; Length 303;
Best Local Similarity 43.6%; Pred. No. 8.4e-36;
Matches 132; Conservative 56; Mismatches 108; Indels 7; Gaps 5;

```

```

QY      4 KIKVVRKRPPLSELEKKKDSIIITVKNCTLYIDPRYKVDMTKYIERHEFIVDKVPDD 63
      4 RLCVCRKRPPLNQREBETTLKDLITVPSDNNVWVHESKQVLDLTRYLQNTQFCFHAFFD 63
DB      4 RLCVCRKRPPLNQREBETTLKDLITVPSDNNVWVHESKQVLDLTRYLQNTQFCFHAFFD 63
QY      64 TVDNFTVYENTIKPLIIDLYENGCVSCFAYGQTSCKTYTMLGS-QPYGSDTPGIFQY 122
      64 KASNELVYQFTAPQPIVESIFRKG-MATCFAYGQTSCKTYTGWGDFSGTAODCSKGIYAL 122
DB      64 KASNELVYQFTAPQPIVESIFRKG-MATCFAYGQTSCKTYTGWGDFSGTAODCSKGIYAL 122
QY      123 AGDIFTEFL-NIYDKNTKGFISFEYIYCGKLYDLOKRKMVALENGKKEVVVVKDKILR 180
      123 VAQDVFLLRNSTYELDKL-VYGFTEYIGAKVYDLOLWKKKLOVLEGNQOIQVAGLQ 181
DB      123 VAQDVFLLRNSTYELDKL-VYGFTEYIGAKVYDLOLWKKKLOVLEGNQOIQVAGLQ 181
QY      181 ILRLVLTKEELIKMDGVLLRKIGVNSONDESSRSHAIINIDKIDINKNTSLGKIAFIDL 240
      182 EKREVCVEEVLNVLVEIGNSCRSTRQTPVVAHSRSHAVQIILK--SGRIMGKTSVLVD 239
DB      182 EKREVCVEEVLNVLVEIGNSCRSTRQTPVVAHSRSHAVQIILK--SGRIMGKTSVLVD 239
QY      241 ASERGADTVSONKQOTQTDGANINRSLALKECIRAMSDSKNHIIPRDSLTLYLADIV 300
      240 AGNERGADTVTKASRRKQLEGAEINKSLALKECIRALDQHTHTPPRASKLTLLVRDSFI 299
DB      240 AGNERGADTVTKASRRKQLEGAEINKSLALKECIRALDQHTHTPPRASKLTLLVRDSFI 299
QY      301 GKS 303
      300 GQN 302
DB      300 GQN 302

```

```

RESULT 22
US-09-722-129-2
; Sequence 2, Application US/09722129
; Patent No. 6582958
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6582958el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1054
; CURRENT APPLICATION NUMBER: US/09/722,129
; CURRENT FILING DATE: 2000-11-24
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Human
US-09-722-129-2

```

```

Query Match      8.6%; Score 589.5; DB 4; Length 492;
Best Local Similarity 43.6%; Pred. No. 1.7e-35;
Matches 132; Conservative 56; Mismatches 108; Indels 7; Gaps 5;

```



```

Db 59 TNNVVRVGRDDE-----VRNSSLAVSTGAMGAEALAIQSPSSMLVTXTY 108
QY 56 IYDKVFDIVDFVNTYENTIKPLIIDLYENGCVSCFAVGOTSGKTYMLGSPYQSGD 115
109 APFKVGPEDQMLPENSVAEMLEQVL-NGYNCTIFAYGOTGTGKTYTMSGD-----LSD 163
QY 116 TPGIFOYAG-----DIFPLNIYDKNTK-GIFISFEYIYCGKLYDL--QKRXVAAL 167
164 SDGILSEGAGLIPRALYQFLPSLDNSNOEYAVKCSYELVNEIRDLVSEELRKPARVF 223
QY 168 EN-----GKEEVVVD-----LKIIRVLTKEBELIKMIDGVLKRGVNSQNDSS 213
224 EDTSRKGNVITIGIESYIKNAGDGLRLR-----EGSHRQVAATKCDLSS 271
QY 214 RSHAILNIDLK-----DINKNTS-----LGKIAFIDLAGSERGADTVSONQ 255
272 RSHSIFITLHRKVVSGMTDETNSLTNNSSDOLPASKLHMVDLAGSENGSGAENR 331
QY 256 TOTDGANINSLALKECIRAMDSDKNHIFPDSSELTQVARDIFVGSKSIMIANISPTI 315
332 ARETGM-INOSLTLTGAVINALVEKAHHIPRESKLTRLQDSLGKTKTSMIVVSSTN 390
QY 316 SCCEOTLNTLRYSRVKFNKSTCINEBDTNTERTST-----LDSKSEMAASSTEN 369
391 TNLREETITLLEYARAKSIRNKP-----QNNQVFRKVLIKDVLDIERLKNIDLNATRKXN 446
QY 370 VV-----IKSNHLSSNNNNKINRGKINDKIERNNI--LKXKSPKPREGFT 414
447 GYVLAESTYKELMDRVQNDLCOEQARKLEVLDIANVKSREQLQVYSKSNQHKKEVEA 506
QY 415 STFGKYSLLNDIDIKKXKKGGLINYKSTLYNDNTINKKNNNNNNNN 462
507 LQQLVNSSTELSVKSENEK-----LKNELVLEIEKKKYEETNEAKITVATDLSQYRE 562
QY 463 -----DNDNNDNNNNNNNDSSKVMNMIMNM--INNINNIN--VNNNNN 506
563 SKEYIASLYEKLDRTERRNKENENNFWNLKFNLLTMRSPHGSFTDETYGFTLLDPA 622
QY 507 NNNNNNNSHNNHLPQPNVAFD--TSPSSLDD-----MNCNLNNDKSIFFLHKXNLRD 559
623 SHELANTHSNOL--LISMTKITEHFQSLDEALOSARSSCAVPNS--SLDLIVSELKD- 676
QY 560 IKLKNSSCDNIMKKKNLHLARHSVGSKLTFMSYDPOKKNKNTFFKSNINMKEDNTPK 619
677 -----SKXSLDLALHSLQDISMSQKLNGISELLELOKDKMESY-----R 719
QY 620 DIYERANVSNNAGNVLLGKNTKTHDITKDNHNDKTNNGVNIINNNSVNSINNSN 679
720 QLVQELRSLYNTL-----QHTHER-----SOKELMYGVNR--DIDALVKT 757
QY 680 MNSINNSNM--NSNSIYKSNVNSNQ-----SISDVQIRVNMEDTSKNKNDIIFDAYS 731
758 TTSLANADILSLYISDQSKFESKQODLIANGKIVSNFLOE-----QNESLYTKA-- 809
QY 732 CDNNMYPNITNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 791
810 -----DLHSHLNTNNSIRKANEI-----MNNRSEEPFLRNA----- 841
QY 792 TSNIQNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 845
842 ASQAEIVGANKERILOKTVENGSQL-----LDSKSKALHSNSRSRMYDCLALAEQKQGVNL 897
QY 846 IIONRDLFEKKKKTTPYNNNN-----IYVNNNMGNNSPRMKYGLCSGHTSIDMKKN 899
898 EYVTLDRLOKVEKHEEDNTKEKQQLDLLESVLGANDN-----LIDSIKTP 945
QY 900 EMKNNEMKONEMKDNHISKNNNS--SSSSSSNNNNIYNNINDDTFQNDYCHANDFTTIR 957
946 HTLEQKITDHLVKGTTSLANHTNELGLGDESLCNLETTIEDTSLVK-----LETTGDRP 1000
QY 958 RKNNTNINSIYNDODIYITINSLANDYMSGTLHFEKTYPTLSTNEDIYAK 1010
1001 SKRELPAATPSWTRDSSLIKETTNLN--LDSKKFVREYETSSNQTNPEPVDYK 1051

```

```

RESULT 25
US-09-723-820-6
; Sequence 6, Application US/09723820
; Patent No. 6468760
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/723,820
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/541,782
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1073
; TYPE: PR1
; ORGANISM: S. pombe
US-09-723-820-6

Query Match          7.9%; Score 538.5; DB 4; Length 1073;
Best Local Similarity 20.7%; Pred. No. 2.9e-31;
Matches 225; Conservative 223; Mismatches 410; Indels 265; Gaps 43;

QY 3 SKIKYVRRKRLPELEKKKKSDIITVKNCTLYID-----EPRIYKVDMTKIERHEF 55
59 TNNVVRVGRDDE-----VRNSSLAVSTGAMGAEALAIQSPSSMLVTXTY 108
QY 56 IYDKVFDIVDFVNTYENTIKPLIIDLYENGCVSCFAVGOTSGKTYMLGSPYQSGD 115
109 APFKVGPEDQMLPENSVAEMLEQVL-NGYNCTIFAYGOTGTGKTYTMSGD-----LSD 163
QY 116 TPGIFOYAG-----DIFPLNIYDKNTK-GIFISFEYIYCGKLYDL--QKRXVAAL 167
164 SDGILSEGAGLIPRALYQFLPSLDNSNOEYAVKCSYELVNEIRDLVSEELRKPARVF 223
QY 168 EN-----GKEEVVVD-----LKIIRVLTKEBELIKMIDGVLKRGVNSQNDSS 213
224 EDTSRKGNVITIGIESYIKNAGDGLRLR-----EGSHRQVAATKCDLSS 271
QY 214 RSHAILNIDLK-----DINKNTS-----LGKIAFIDLAGSERGADTVSONQ 255
272 RSHSIFITLHRKVVSGMTDETNSLTNNSSDOLPASKLHMVDLAGSENGSGAENR 331
QY 256 TOTDGANINSLALKECIRAMDSDKNHIFPDSSELTQVARDIFVGSKSIMIANISPTI 315
332 ARETGM-INOSLTLTGAVINALVEKAHHIPRESKLTRLQDSLGKTKTSMIVVSSTN 390
QY 316 SCCEOTLNTLRYSRVKFNKSTCINEBDTNTERTST-----LDSKSEMAASSTEN 369
391 TNLREETITLLEYARAKSIRNKP-----QNNQVFRKVLIKDVLDIERLKNIDLNATRKXN 446
QY 370 VV-----IKSNHLSSNNNNKINRGKINDKIERNNI--LKXKSPKPREGFT 414
447 GYVLAESTYKELMDRVQNDLCOEQARKLEVLDIANVKSREQLQVYSKSNQHKKEVEA 506
QY 415 STFGKYSLLNDIDIKKXKKGGLINYKSTLYNDNTINKKNNNNNNNN 462
507 LQQLVNSSTELSVKSENEK-----LKNELVLEIEKKKYEETNEAKITVATDLSQYRE 562
QY 463 -----DNDNNDNNNNNNNDSSKVMNMIMNM--INNINNIN--VNNNNN 506
563 SKEYIASLYEKLDRTERRNKENENNFWNLKFNLLTMRSPHGSFTDETYGFTLLDPA 622
QY 507 NNNNNNNSHNNHLPQPNVAFD--TSPSSLDD-----MNCNLNNDKSIFFLHKXNLRD 622
623 SHELANTHSNOL--LISMTKITEHFQSLDEALOSARSSCAVPNS--SLDLIVSELKD- 676
QY 560 IKLKNSSCDNIMKKKNLHLARHSVGSKLTFMSYDPOKKNKNTFFKSNINMKEDNTPK 619

```


Db 677 -----SKSLDADALEHSIADIDISMSQKXGNGISSSELIELQXDMKSY-----R 719
Qy 620 DILYESNVSNMNGNVLGLKNTKTHDIDSTYDENHNDKINNGVYININNSNVNSINN 679
Db 720 OLVELRSLVYL-----QHTHEE-----SQKELMYGVN-----DIDALVYTC 757
Qy 680 MNSINNSNM--NNSIYKSYNSNO-----SIDVOIRYVEMDTSKNNNDNIFPAIS 731
Db 758 TTSINDADILISDYISQKSFESKQODLANIGKISNFIQF-----QESLYTKA-- 809
Qy 732 CDNNMYNITNNNNNNNNNNNNNNNNIDVENYNNRDGTTNSMKLYAVNSHLFOPDNKN 791
Db 810 -----DILHSLDNTNSNIRKANEL-----MNNSEFLRNA----- 841
Qy 792 TSNIGNNTKNNODGAVNVSMAFCNYLNDKATLIDANKQXDKIHGCD-----NN 845
Db 842 ASQAEIVGANKERIOKTVENGSOUL-----LDSKSKAISHNSRSMYDHCIALAESOKQGVNL 897
Qy 846 IIONRNDPEKKKKTNPVNNNN-----IVYNNMGNNSNPRMKYGLGSHTSIDNMKN 899
Db 898 EVOTLDRLLQVKYKHESEDNTEKHOQLDLLESILVGNNDN-----LIDSTKTP 945
Qy 900 EMKNNEMKDNEMKDNHIKSNNNNS--SSSSSSNNNINNNINDPTFQNDYCHNDNTFTIR 957
Db 946 HTELQKITDHYLKGTSLANTNTELGLGDESLCNLETTIEDISLVK-----LETTGOTP 1000
Qy 958 RKNNTNINSNIYQNDIITYTINSINDYMSNTLLHFEKTYPTLSTNEDIYNK 1010
Db 1001 SKRELPAFPWTRDSLSLIKETTINL--LDSDKKFEVRETYTSSNOTNEPDVYDK 1051

RESULT 26
US-09-417-485D-6
Sequence 6, Application US/09417485D
Patent No. 6541202
GENERAL INFORMATION:
APPLICANT: Metz, David M.
APPLICANT: Metz, Anneke M.
TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes
FILE REFERENCE: 47714-5009-US
CURRENT APPLICATION NUMBER: US/09/417, 485D
CURRENT FILING DATE: 2002-06-14
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2184
TYPE: PRT
ORGANISM: Plasmodium falciparum
FEATURE:
NAME/KEY: unsure
LOCATION: (330). (335)
OTHER INFORMATION: Xaa at position 330 = Leu or Ile;
OTHER INFORMATION: Xaa at position 335 = Asp or Gly.
US-09-417-485D-6

Query Match 7.9%; Score 538.5; DB 4; Length 2184;
Best local similarity 21.4%; Pred. No. 8e-31;
Matches 330; Conservative 234; Mismatches 489; Indels 487; Gaps 73;

Qy 4 KIKVAVRRPLSELEKKKSDIITVKNCTLY-----IDERYKQDMKTYIERH--EPI- 56
Db 429 QIKCTKNDISHIITSRKENHLFHVQKLENNYKAPNINKQARKTKILKYVNYFKEFLN 488
Qy 57 -----VDKVF-----DQVNFVYENTIKPLIIDYE 84
Db 469 NVINTKFGKIYKFFPRKHILANKIKHIFILQIILKYYRIINIMNKRFIQKYVDFPF 548
Qy 85 NGCVSCRAYGQTGSGKTYTMLGSOPIYGSOTPGIFQVAAADIPTFLNLYDKONTKGLFI 144
Db 549 KNYDFLSFSF-----KTYKINFVYITKCKIPILGSKNPF--KIFLKNVKKVFLLF 599

Qy 145 SFYEIYC-----GKLYDLLQKRNVAALENKGEVNVKDL-----KIILVLTKEEL 190
Db 600 NYKESSTINGVMKNIIVKNIIFQKK-----ISKYINIKNILLKNIFDNNYENKILH--RNNKEI 655
Qy 191 ILKMDIGVLARKIGVNSQNDSSRSIALNIDLKINKTSLQKIAFIDLAGEBGADT- 249
Db 656 ITVININIKI-----YKKKNDNLNNSFKIKYTTLFNKLRRK-YFNKIKKINIAIDQRHLMNR 710
Qy 250 -----VSOKOTQTOGANINRSLLALEGICRAMSDKN-----HIPF 286
Db 711 LIYFLENYFIMPLIRREFFLTKSEOT-----LHKTIFFPRKIMNFTKISNCLVHOIF 764
Qy 287 RDESLTK-----VLADIFVQKSKSIMANISPTISCCBOTLNTLRYSRVNPKXKST 339
Db 765 RNNKLLKRNPRKPDYQNMENYVKKG-----EKIKTKYIP-IKMKKKST 809
Qy 340 --CINEEDNTERTISLDSKSEMAASSIENVYKSNHLSNNNNKINRG----- 389
Db 810 NKCINNK--FSKKCIPKKKKKULYITRANNIFIKKDEKSKRTNLLNKSIDNLYK 866
Qy 390 KINDKI-----ERNNIIKNSF-----DK 408
Db 867 EINKSVRYPIYKFFYIKKIFALKKMYTHMMMAEKSNIKLEAFKGFIFPAQEKH 926
Qy 409 PREGFTSTFGKYSLND-----IDKIK-----XNKK--GLINYKSTLYND 447
Db 927 ILKFSHFQONKINVGKFNKLIRIKNIIKONGSIGYKNDKFTLHIIKNSKNKN 986
Qy 448 NTINKKNNNNNNNDNDNDNDNDNDNDNDNDSSSMYNNMIMHMINNNINNNVNNNNN 507
Db 987 NKKKNKNNYNN 1036
Qy 508 NNNNNNNHNNHLPQVNAFTDTSDFSLDDMNCHLNNDKSIPLHKKLNLDNITKLKRS 567
Db 1037 KKAQNNKKN-----IDSN--LEKKKKITYY-----KIKKIIE 1069
Qy 568 CDNIMNKKNNLHLARHVSGLTMFSYDQKKNKNTFFKSNINKMEDTPK-----D 620
Db 1070 KRNFMKL-----LNSINFIQKLR--NWIPKKG--LRPLNLTSLNPEIYQRIFE 1120
Qy 621 ILVESRNVSMNGNVLGLKNTKTHDIDSTYDENHNDKINNGVYININNSNVNSINN 680
Db 1121 ILKSKSEFVFHIIANNLER-----EKDKONIKKKKKYNN--KNENPVSNNINCNPSL 1172
Qy 681 NSINNSMNSIYKSYNSNQSISDYOI-----RYVNEMDTSKND 723
Db 1173 KCLGNMHRNNNSLFK--NTLTGTGETELKQKMLHYLKWYFKKMKKTYIKKYLKONK 1229
Qy 724 NIF-FDAISCDNNMYPNITNNNNNN--NNNNNNNNIDVENYNNRDGTTNSMKLY 776
Db 1230 KIYAVICIGDPSNCYEHI-----NHNYLFKILKVFNNINNFETILFKR-----SPFLY 1279
Qy 777 AYNSHNLF--QPNNTK--TSNIQNTN--TNKNQOGVNVSMN--FCHYNLDKRYLI 827
Db 1280 NKNLNSFLSYPVNVKVSFGIYLRNLRRLIKSHLDNHHFLLNOMFKTKSKSDLYIFA 1339
Qy 828 DLNKEQKDKK-----IHGCDNIIIONRNDPEKKKKTNYF-----NNNNIYVNV 871
Db 1340 DSYSLQVQKDDIMTITTYIRYYLYNYSIKFELKRNKQIFFOQFQENQMGVLYSV 1399
Qy 872 -----NNMGNNSNPRMKYGLGSHTSIDNMKNEMKDNEMKDNHNIK 917
Db 1400 RDKKRVENIKKMYLNSMKIINHDEILSLKSSININN-----KNFMICTNHEQTEEK 1453
Qy 918 SNNNSSSSSSNNNINNNINDDTFQNDYCHNDNTFTIRKNTNINSNIYQNDIITY 977
Db 1454 GNTQNKKEGDIYIGPIYNSPSTY--TTHSSNNY--KGNNIHVSQD--YKNDGLHK 1505
Qy 978 INSLNDYMSNTLLHFEKTYPTLSTNEDIYNKMEGCHIRLDDQKYDDNDNNNVNNDNN 1037
Db 1506 GNN-----SNMB-CTVQK--C-----CANNNNNNNNNN 1531
Qy 1038 KNNVDNNVNNVNDNNVNDNNKNNVNDNNVNDNNDDVDVFNHINKNNNEKXLYSPQKQVDT 1097


```

; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 967
; TYPE: PRF
; ORGANISM: Loligo pealeii
US-09-914-259-21

```

```

Query Match      7.3%; Score 500.5; DB 4; Length 967;
Best Local Similarity 22.1%; Pred. No. 1,6e-28;
Matches 217; Conservative 189; Mismatches 403; Indels 175; Gaps 35;

```

```

QY 5 IKVVRKRPLSELEKKKSDITVKNCTLYIDPRYKVDTKYIERHEFIVDKYFDDT 64
DB 9 IKVCHVRPLNEAEERAGSKFIL-----KFPDDDSISAGKVFVDPKYLKPN 55
QY 65 VDNFTYENTIKLLIIDYENGVCSCFAYGOTSGKTYTMLGSPYGGSDTP---GIRQ 121
DB 56 VSGEYVYVNGAKPIVADVL--SGCNGTIFAYGOTSSGKTHM---EGVLDRKSMHGILP 109
QY 122 YAAGDIFPLNYDKONTKGIFFSFEIYCGKLYDLLQKKQVVALENGKEVAVVDLKI 181
DB 110 RIVQDIFNYTIGMDENLEPHIKSYEYIYDKRIDLDVTKTMLAVHEDKNRPFPVKAT 169
QY 182 LRVLYKEELIKMID-GVLLRKIGVNSQNDSSRSHAILNIDKIDINKTS---LQKIAF 237
DB 170 ERFVSSPEEVEVIDEGKKNRHVAVTNMNHSRSHSVFLINQENVEVETOKLSGLYL 229
QY 238 IULAGERADYVSONKQOTODGANINRSLAKKECIKRAM-BDDKNHIPPDELTQVLR 296
DB 230 VDLAGEKYSKTAEG-AVLDEKKNINKSLGALGNVISALADKNKHVPYRDKLRIQ 288
QY 297 DIFVGSCKIMIANISPTISCEQTLNLTLYSRVKNFKKSTCINBE-----DDT 347
DB 289 ESIQGNARTTIVCCSPASINSESTSLIFGQRAITIK-VYSVBEELTADMKRKYER 347
QY 348 NTERISILDSKSGEMNASSIENVIKSNHLLSNNNNNKINRGKINDI---ERNNILLK 404
DB 348 EKERVYTKLAKTAKLA-----ELQRMRTGQAVSVEQVLDKEDV 387
QY 405 SPFKPREGFTS-TFGKYSLNDIDKIKAKKGLINYKSLTYNDNTINK----- 453
DB 388 PAFSPATSTTSLAGLIASNNEGDRTOLEBER-LKLYOQLDDXODEINNQSOLIEKLK 446
QY 454 -----HNNNNNNNNNDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 504
DB 447 MNEOEELIAOSRDYENLOQDMSRIQADNESAKOBEYKVLQALEELAMNYDQSQEYEDK 506
QY 505 NNNNNNNNNSHNNHLPQNYAFTDTSDFSSLDMMNGCHLNNDKSIPLHKKNLDNT--KL 562
DB 507 NKENENLSEELNQL-----STLNSLQWBLDQLKDSMHHKRVAVDMINLL 553
QY 563 KRRSSCDNIMKKKNNLHLARHSVSKLTFSTYDPOKKNONTFFKSNINKMEONTKOLL 622
DB 554 KDLGIDIGTIVG--GNAAEFKPTAGS-----GEKIEEFYARLYISKMKKS----- 596
QY 623 YESRNVSNNGVNLGLKNKTHHDISTJQENHNDKINNGVINIINNSVNSINNSNMS 682
DB 597 -EYKTVSRNNQI-----EYTOODNFKKLTETHEKDSLNCGLLQOEBAKASLQEAIKDS 650
QY 683 INNSNNNSIYKSNYSNQSISDVOIRYV---NEMDTSNKNNDNIFPDALSCDNNMPY 738
DB 651 ENKCKMLIEDVDSLNEBYAKLKAQOEQNHLLALESEKRETSQASETR---EVLKQEMHR 707
QY 739 NITNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 795
DB 708 E-OHQKOLSLRDEISEKATVNLKD-DNQRSLALEKLQD--DYDKLQOEVEYRAKL 763

```

```

QY 796 QNINIRKNNQ-----DGNVNSKNFCHYINIDKNVILIDLNK-----EQDK 837
DB 764 ADLSIQIDRREQAKODLKGLFEYFAKELQTLH-NLR-KLFLVODLQKVKKSCSKTEBEDE 821
QY 838 NIHGCNNIIONRNDPEKKKTTFYNN--NIYVNNNNNNNS-----PRKXYGLGCS 889
DB 822 DTGG-----NAAQOKKISFLENNLEQLTKYKQIVRNADLRCELPLKLEKL-- 868
QY 890 HTSDNKNKNEMKNEMKNEMKD 913
DB 869 RATMERVKSLESALKAKGAMRD 892

```

```

RESULT 30
US-09-541-782-2
Sequence 2, Application US/09541782
Patent No. 6284480

```

```

; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Berand, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/541,782
; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1184
; TYPE: PRF
; ORGANISM: A. nidulans
US-09-541-782-2

```

```

Query Match      7.3%; Score 500; DB 3; Length 1184;
Best Local Similarity 21.1%; Pred. No. 2,4e-28;
Matches 253; Conservative 194; Mismatches 415; Indels 338; Gaps 48;

```

```

QY 2 NSKIKVVRKRPLSELEKKKSDITVKNCTLYIDPRYKVDM-TKYIERHEFIVDKY 60
DB 79 DSIHVVACRGKRENEVEKNGSVLQTEG-----VKGTVELSMGPNAVSNKTYTFDKY 133
QY 61 FDDTVDNFTYENTIKLLIIDYENGVCSCFAYGOTSGKTYTMLGSPYGGSDTPGIF 120
DB 134 FSAADQIIVYEDVVLPIYEMLA-GYNCTIFAYGOTGKTYTMSGD---MTDTLGL 188
QY 121 QYRAGDI-----FTPLNYDKONTKGIFFSFEIYCGKLYDL-----QKR 161
DB 189 SDNAGIIPVLYSLFAKLADTEST--VKCSFTELVNEELRDLISAEENPKLIYDNEOKK 246
QY 162 KVAALENGKKEVVVD-----KLIRVLTKBELIKMIDGVLLRKIGVNSQNDSSRSHA 217
DB 247 GHVSTIVQWMEETIYISATAGIKYLD-----QSHKROYAATKCDLSRSRHT 294
QY 218 ILNIDKIDINKTSL-----GKIAFIDLASERGADTVSONKQOTODGANINRSLAL 270
DB 295 VFTITV-NIKRTSESGEYVCGKMLVLDLASENIGRGAERKRA-TEAGLINSLLTL 352
QY 271 KECIRAMSDKNHIPPDELTQVLDIFVGSKSMIANISPTISCEQTLNLTAKYSSR 330
DB 353 GRVYNALVRSQHIPIYRESKLTRELDSLGGRTKCTIITATMSPARSNLEETISTLDYAFR 412
QY 331 VKNFKKSTCIN-----EEDDTMERISILDSKSGEMNASSIENVVI--KS 374
DB 413 AKNIIRKPO-INSTMEKMTLLEFTEIEKLAELIATHRNGVNVSVSEYEMKJENES 471
QY 375 NHLSSNNNNKIN--RGKINDKIERNNILKNKSPDKRREGFTSTFGKYSLNDI---DK 428
DB 472 BRIISEGQAKIESMESSLRHVQBELTLTSK-FNDLKKDNDPTLALCSTNDVLOQTOI 530
QY 429 IKKKK-----KGLIN-YKSTLYNDNTINKKNNNNNNNNND 463
DB 531 VLONTBAOLEEBEMLRCAHEETEHOLODVQKGLISTLQGTVEDINLSQK-----LD 582

```


THIS PAGE BLANK (10/10)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 2, 2003, 16:20:46 ; Search time 85 Seconds

(without alignments)
2397.392 Million cell updates/sec

Title: US-10-006-780-2

Perfect score: 6842
Sequence: 1 MNSKIKVVRKRPSELEK.....KXIVQDINKSMDDNNHKK 1288

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:
1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	6842	100.0	1288	US-10-006-780-2	Sequence 2, Appli
2	1815	26.5	355	US-10-006-780-6	Sequence 6, Appli
3	1703	24.9	332	US-10-006-780-10	Sequence 10, Appli
4	1702	24.9	332	US-10-006-780-4	Sequence 8, Appli
5	1692	24.7	361	US-10-006-780-8	Sequence 4, Appli
6	768.5	11.2	363	US-10-159-151-10	Sequence 10, Appli
7	768.5	11.2	381	US-10-159-151-6	Sequence 6, Appli
8	768.5	11.2	1368	US-10-159-151-2	Sequence 2, Appli
9	758.5	11.1	343	US-10-159-151-8	Sequence 8, Appli
10	758.5	11.1	373	US-10-159-151-4	Sequence 4, Appli
11	750.5	11.0	725	US-09-849-602-20	Sequence 20, Appli
12	696.5	10.2	319	US-09-723-276-2	Sequence 2, Appli
13	696	10.2	807	US-09-820-843A-108	Sequence 108, App
14	694.5	10.2	861	US-09-820-843A-109	Sequence 109, App
15	676	9.9	2150	US-10-135-322-17	Sequence 17, Appli

16	672	9.8	410	10	US-09-925-300-1074	Sequence 1074, Ap
17	596	8.7	1331	15	US-10-087-464-50	Sequence 50, Appli
18	589.5	8.6	303	12	US-10-422-866-4	Sequence 4, Appli
19	589.5	8.6	492	12	US-10-422-866-2	Sequence 2, Appli
20	538.5	7.9	2184	12	US-10-304-095-6	Sequence 6, Appli
21	518.5	7.6	1610	15	US-10-155-533-9	Sequence 9, Appli
22	491	7.2	464	12	US-10-161-051-119	Sequence 119, Appli
23	460.5	6.7	1388	15	US-10-146-473-82	Sequence 82, Appli
24	454.5	6.6	1269	12	US-10-342-224-62	Sequence 62, Appli
25	453	6.6	383	10	US-09-883-096-5	Sequence 5, Appli
26	451	6.6	864	10	US-09-883-096-2	Sequence 2, Appli
27	446	6.5	1648	15	US-10-157-031-38	Sequence 38, Appli
28	429	6.3	338	10	US-09-883-096-4	Sequence 4, Appli
29	395.5	5.8	1103	10	US-09-847-874A-1	Sequence 1, Appli
30	388.5	5.7	596	12	US-10-167-831-2	Sequence 2, Appli
31	376.5	5.5	1253	12	US-10-363-798-2	Sequence 2, Appli
32	374.5	5.5	794	12	US-10-304-095-8	Sequence 8, Appli
33	373	5.5	1191	15	US-10-087-464-47	Sequence 47, Appli
34	348.5	5.1	665	11	US-09-820-843A-107	Sequence 107, App
35	342.5	5.0	6281	9	US-09-815-242-12996	Sequence 12996, A
36	336.5	4.9	460	10	US-09-925-300-1228	Sequence 1228, Ap
37	333.5	4.9	890	15	US-10-171-311-196	Sequence 196, App
38	333	4.9	652	9	US-09-351-794A-2	Sequence 2, Appli
39	327	4.8	1093	10	US-09-801-368-392	Sequence 392, App
40	325.5	4.8	959	12	US-10-342-224-40	Sequence 40, Appli
41	323.5	4.7	2086	9	US-09-815-242-5639	Sequence 5639, Ap
42	323.5	4.7	5795	9	US-09-815-242-12610	Sequence 12610, A
43	316.5	4.6	887	12	US-09-948-029-130	Sequence 130, App
44	314.5	4.6	1349	9	US-09-815-242-5898	Sequence 5898, Ap
45	314.5	4.6	1349	9	US-09-815-242-13137	Sequence 13137, A

ALIGNMENTS

RESULT 1
US-10-006-780-2
; Sequence 2, Application US/10006780
; Publication No. US20030104496A1
; GENERAL INFORMATION:
; APPLICANT: Sakowitz, Roman
; APPLICANT: Beraud, Christophe
; APPLICANT: Guo, Jun
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND
; FILE REFERENCE: CYTOPO83
; CURRENT APPLICATION NUMBER: US/10/006,780
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1288
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-006-780-2

Query Match	100.0%;	Score 6842;	DB 15;	Length 1288;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1288;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MNSKIKVVRKRPSELEK	DDSDIIVKNNCTIYIDPRYKVDMTKYEHEFI	VDVY 60
DB	1	MNSKIKVVRKRPSELEK	DDSDIIVKNNCTIYIDPRYKVDMTKYEHEFI	VDVY 60
QY	61	PDDYDNFVYENTIKPLIIDYENGCVSCFAYQGTGSGKTYT	MLGSPYQSDTPGIF	120
DB	61	PDDYDNFVYENTIKPLIIDYENGCVSCFAYQGTGSGKTYT	MLGSPYQSDTPGIF	120
QY	121	QYAAGDIFFLNIYKDNKKGIFISFYERYCGKLDLLOKRMVNALENGKEV	YVVKDK 180	
DB	121	QYAAGDIFFLNIYKDNKKGIFISFYERYCGKLDLLOKRMVNALENGKEV	YVVKDK 180	

[illegible]

```

Db      1261  NGDIILLANKKLVQDNINKSNMDHNHKK 1288

RESULT 2
US-10-006-780-6
; Sequence 6, Application US/10006780
; Publication No. US20030104496A1
; GENERAL INFORMATION:
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; APPLICANT: Guo, Jun
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND
; TITLE OF INVENTION: METHODS FOR ITS USE
; FILE REFERENCE: CYTO083
; CURRENT APPLICATION NUMBER: US/10/006,780
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 355
; TYPE: PRT
; ORGANISM: P. Falciparum
US-10-006-780-6

Query Match      26.5%; Score 1815; DB 15; Length 355;
Best Local Similarity 100.0%; Pred. No. 5,6e-120; Indels 0; Gaps 0;
Matches 354; Conservative 0; Mismatches 0;

QY      4      KIKVVRKRPSELEKKKKKSDITVKNKCTLYIDEPRYKVDMTKRIERHEFIVDKVPDD 63
        |||
        2      KIKVVRKRPSELEKKKKSDITVKNKCTLYIDEPRYKVDMTKRIERHEFIVDKVPDD 61
        |||
        64      TVDNFTYVENTIKPLIIDLYENGCVCSGFAYGQTSGKTYTMLGSPYGSQDTPGIFQYA 123
        |||
        62      TVDNFTYVENTIKPLIIDLYENGCVCSGFAYGQTSGKTYTMLGSPYGSQDTPGIFQYA 121
        |||
QY      124      AGDIFTFPLNYDKONTKCIFISFEIYTGKLYDLQKRWAAALENGKEVVDKILR 183
        |||
        122      AGDIFTFPLNYDKONTKCIFISFEIYTGKLYDLQKRWAAALENGKEVVDKILR 181
        |||
QY      184      VLTEBELILKKIDGVLTKIGVSNONDESSSHAILNIDLDKINKTSLGKIAFIDLGS 243
        |||
        182      VLTEBELILKKIDGVLTKIGVSNONDESSSHAILNIDLDKINKTSLGKIAFIDLGS 241
        |||
Db      244      ERGADTVSONKQOTDGCANINRSILALKECIRAMDSDKNH1PFRSELTLYLRDIFVGS 303
        |||
        242      ERGADTVSONKQOTDGCANINRSILALKECIRAMDSDKNH1PFRSELTLYLRDIFVGS 301
        |||
QY      304      KSIMIANISPTISCEQTLNLTLYRSSRVNFRANKSTCINEDBDNTERISILDS 357
        |||
        302      KSIMIANISPTISCEQTLNLTLYRSSRVNFRANKSTCINEDBDNTERISILDS 355
        |||

RESULT 3
US-10-006-780-10
; Sequence 10, Application US/10006780
; Publication No. US20030104496A1
; GENERAL INFORMATION:
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; APPLICANT: Guo, Jun
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND
; TITLE OF INVENTION: METHODS FOR ITS USE
; FILE REFERENCE: CYTO083
; CURRENT APPLICATION NUMBER: US/10/006,780
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 332

```


TYPE: PRT
ORGANISM: P. falciparum
US-10-006-780-10

Query Match 24.9%; Score 1703; DB 15; Length 332;
Best Local Similarity 100.0%; Pred. No. 4e-112;
Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 IKVAVRKRLPSELEKKKSDIITVKNCTIYIDEPYKVDMTKYIERHEFIYDKVEDDT 64
DB 1 IKVAVRKRLPSELEKKKSDIITVKNCTIYIDEPYKVDMTKYIERHEFIYDKVEDDT 60
QY 65 VDNFTYVENTIKPLIIDLYENGCVSCFAYGQTSGKTYTMLGSPYQSDPTGFIFOYA 124
DB 61 VDNFTYVENTIKPLIIDLYENGCVSCFAYGQTSGKTYTMLGSPYQSDPTGFIFOYA 120
QY 125 GDIFFELNLYDKNTKGFISFEIYCGKLYDLQKRWVAALENGKEVYVVDKILRV 184
DB 121 GDIFFELNLYDKNTKGFISFEIYCGKLYDLQKRWVAALENGKEVYVVDKILRV 180
QY 185 LTKEELILKMDGVLRLKIGVNSQNDSSRSHAILNIDLKDKINNTSLGKIAFIDLAGE 244
DB 181 LTKEELILKMDGVLRLKIGVNSQNDSSRSHAILNIDLKDKINNTSLGKIAFIDLAGE 240
QY 245 RGADTVSQRKQOTQDGANINRSLLALKECIRAMSDKNHIFRDSSELTQVLRDIFVGSK 304
DB 241 RGADTVSQRKQOTQDGANINRSLLALKECIRAMSDKNHIFRDSSELTQVLRDIFVGSK 300
QY 305 SIMIANISPTISCEQTLNLTIRYSSRVKNFKN 336
DB 301 SIMIANISPTISCEQTLNLTIRYSSRVKNFKN 332

RESULT 4

US-10-006-780-4
Sequence 4, Application US/10006780
Publication No. US20030104496A1
GENERAL INFORMATION:
APPLICANT: Sakowicz, Roman
APPLICANT: Beraud, Christophe
APPLICANT: Guo, Jun
APPLICANT: Freedman, Richard
TITLE OF INVENTION: NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND
FILE REFERENCE: CYTOPO83
CURRENT APPLICATION NUMBER: US/10/006,780
CURRENT FILING DATE: 2001-11-30
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 332
TYPE: PRT
ORGANISM: Plasmodium falciparum
US-10-006-780-4

Query Match 24.9%; Score 1702; DB 15; Length 332;
Best Local Similarity 100.0%; Pred. No. 4.7e-112;
Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KIKVVRKRLPSELEKKKSDIITVKNCTIYIDEPYKVDMTKYIERHEFIYDKVEDDT 63
DB 1 KIKVVRKRLPSELEKKKSDIITVKNCTIYIDEPYKVDMTKYIERHEFIYDKVEDDT 60
QY 64 TVDNFTYVENTIKPLIIDLYENGCVSCFAYGQTSGKTYTMLGSPYQSDPTGFIFOYA 123
DB 61 TVDNFTYVENTIKPLIIDLYENGCVSCFAYGQTSGKTYTMLGSPYQSDPTGFIFOYA 120
QY 124 AGDIFFLNIYDKNTKGFISFEIYCGKLYDLQKRWVAALENGKEVYVVDKILRV 183
DB 121 AGDIFFLNIYDKNTKGFISFEIYCGKLYDLQKRWVAALENGKEVYVVDKILRV 180
QY 184 VLTKEELILKMDGVLRLKIGVNSQNDSSRSHAILNIDLKDKINNTSLGKIAFIDLAGE 243
DB 181 VLTKEELILKMDGVLRLKIGVNSQNDSSRSHAILNIDLKDKINNTSLGKIAFIDLAGE 240

DB 181 VLTKEELILKMDGVLRLKIGVNSQNDSSRSHAILNIDLKDKINNTSLGKIAFIDLAGE 240
QY 244 ERGADTVSQRKQOTQDGANINRSLLALKECIRAMSDKNHIFRDSSELTQVLRDIFVGSK 303
DB 241 ERGADTVSQRKQOTQDGANINRSLLALKECIRAMSDKNHIFRDSSELTQVLRDIFVGSK 300
QY 304 KSIMIANISPTISCEQTLNLTIRYSSRVKNFKN 335
DB 301 KSIMIANISPTISCEQTLNLTIRYSSRVKNFKN 332

RESULT 5

US-10-006-780-8
Sequence 8, Application US/10006780
Publication No. US20030104496A1
GENERAL INFORMATION:
APPLICANT: Sakowicz, Roman
APPLICANT: Beraud, Christophe
APPLICANT: Guo, Jun
APPLICANT: Freedman, Richard
TITLE OF INVENTION: NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND
FILE REFERENCE: CYTOPO83
CURRENT APPLICATION NUMBER: US/10/006,780
CURRENT FILING DATE: 2001-11-30
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 361
TYPE: PRT
ORGANISM: P. falciparum
US-10-006-780-8

Query Match 24.7%; Score 1692; DB 15; Length 361;
Best Local Similarity 99.4%; Pred. No. 2.7e-111;
Matches 331; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 KIKVVRKRLPSELEKKKSDIITVKNCTIYIDEPYKVDMTKYIERHEFIYDKVEDDT 63
DB 2 KIKVVRKRLPSELEKKKSDIITVKNCTIYIDEPYKVDMTKYIERHEFIYDKVEDDT 61
QY 64 TVDNFTYVENTIKPLIIDLYENGCVSCFAYGQTSGKTYTMLGSPYQSDPTGFIFOYA 123
DB 62 TVDNFTYVENTIKPLIIDLYENGCVSCFAYGQTSGKTYTMLGSPYQSDPTGFIFOYA 121
QY 124 AGDIFFLNIYDKNTKGFISFEIYCGKLYDLQKRWVAALENGKEVYVVDKILRV 183
DB 122 AGDIFFLNIYDKNTKGFISFEIYCGKLYDLQKRWVAALENGKEVYVVDKILRV 181
QY 184 VLTKEELILKMDGVLRLKIGVNSQNDSSRSHAILNIDLKDKINNTSLGKIAFIDLAGE 243
DB 182 VLTKEELILKMDGVLRLKIGVNSQNDSSRSHAILNIDLKDKINNTSLGKIAFIDLAGE 241
QY 244 ERGADTVSQRKQOTQDGANINRSLLALKECIRAMSDKNHIFRDSSELTQVLRDIFVGSK 303
DB 242 ERGADTVSQRKQOTQDGANINRSLLALKECIRAMSDKNHIFRDSSELTQVLRDIFVGSK 301
QY 304 KSIMIANISPTISCEQTLNLTIRYSSRVKNFKN 336
DB 302 KSIMIANISPTISCEQTLNLTIRYSSRVKNFKN 334

RESULT 6

US-10-159-151-10
Sequence 10, Application US/10159151
Publication No. US20030036075A1
GENERAL INFORMATION:
APPLICANT: CytoKinetix, Inc.
APPLICANT: Beraud, Christophe
APPLICANT: Guo, Jun
APPLICANT: Freedman, Richard
APPLICANT: Patel, Umesh A.
APPLICANT: Davies, Katherine A.


```

Db      283 ACTNODVVMKTHPLIOHIF-NGSNATCFAYGOTGAGKTYMTGTH-----ENPGLYALA 336
      124 AGDIFFLINIDYDNDKMGIFISFEYIYCGLYDLQKRKVAALENKEKEVVKDKILR 183
      337 AKDIFRQLEVSQPRKHLFWMISFEYIYCGLYDLNRRKRLPAREDSKHMVQIVGLQELQ 396
Qy      184 VLTKEELILKMDIGVLLRKIGVNSONDESSRSHAILNIDKIDINKNTSLGKIAFIDLGS 243
      397 VDSVELLEVLILKSGKSRSTGATGVNADSSRSHAVIQIQKDSAKRT-FGRISFIDLGS 455
Db      244 ERADTVVSQNKQOTQDGANINRSLALKECIRAMSDSKNHI PRDSELTQVLRDIFVGS 303
      456 ERAADARDSRQRTKMEGAEINOSLALKECIRALDQEHHTPPROSKLTQVLRDIFVGS 515
Qy      304 KSIIMINISPTISCCQOTLNTLRYSSRVKFKKSTCINEDDTN 348
      516 KTCMIANISPSHVATEHTLNTLRYADRVKELKKGKICTSVTSRN 560
```

RESULT 9

```

US-10-159-151-8
; Sequence 8, Application US/10159151
; Publication No. US20030036075A1
; GENERAL INFORMATION:
; APPLICANT: CytoKinetics, Inc.
; APPLICANT: Beraud, Christopher
; APPLICANT: Guo, Jun
; APPLICANT: Freedman, Richard
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: KINI-3 MOTOR PROTEIN AND METHODS FOR ITS USE
; FILE REFERENCE: 020552-002400US
; CURRENT APPLICATION NUMBER: US/10/159,151
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US/09/967,908
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 09/675,227
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide
; OTHER INFORMATION: Fragment E213-S546.
US-10-159-151-8
```

```

Query Match      11.1%; Score 758.5; DB 15; Length 343;
Best Local Similarity 48.9%; Pred. No. 1.5e-45;
Matches 161; Conservative 58; Mismatches 103; Indels 7; Gaps 3;

Qy      4 KIRVVRRKRPPLSELEKKKDSIIIVKNNCTLYIDEPKYVDMTKYIERHEFIVDKVFD 63
      12 KIRVCVRKRPPLGREGVRRGEINIIIVEDKETLLVHEKKEAVDLQYIILGHVFFDEVFGE 71
Db      64 TVDNFVYENTIKPLIILIDYENGCVSCFAYGOTGSGKTYTLMGSOPYGOSDTPGIFQYA 123
      72 ACTNODVVMKTHPLIOHIF-NGSNATCFAYGOTGAGKTYMTGTH-----ENPGLYALA 125
Qy      124 AGDIFFLINIDYDNDKMGIFISFEYIYCGLYDLQKRKVAALENKEKEVVKDKILR 183
      126 AKDIFRQLEVSQPRKHLFWMISFEYIYCGLYDLNRRKRLPAREDSKHMVQIVGLQELQ 185
Db      184 VLTKEELILKMDIGVLLRKIGVNSONDESSRSHAILNIDKIDINKNTSLGKIAFIDLGS 243
      186 VDSVELLEVLILKSGKSRSTGATGVNADSSRSHAVIQIQKDSAKRT-FGRISFIDLGS 244
Qy      244 ERADTVVSQNKQOTQDGANINRSLALKECIRAMSDSKNHI PRDSELTQVLRDIFVGS 303
      245 ERAADARDSRQRTKMEGAEINOSLALKECIRALDQEHHTPPROSKLTQVLRDIFVGS 304
```

```

Qy      304 KSIIMINISPTISCCQOTLNTLRYSSRVK 332
      305 KTCMIANISPSHVATEHTLNTLRYADRVK 333
Db
```

RESULT 10.

```

US-10-159-151-4
; Sequence 4, Application US/10159151
; Publication No. US20030036075A1
; GENERAL INFORMATION:
; APPLICANT: CytoKinetics, Inc.
; APPLICANT: Beraud, Christopher
; APPLICANT: Guo, Jun
; APPLICANT: Freedman, Richard
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: KINI-3 MOTOR PROTEIN AND METHODS FOR ITS USE
; FILE REFERENCE: 020552-002400US
; CURRENT APPLICATION NUMBER: US/10/159,151
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US/09/967,908
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 09/675,227
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fragment
; OTHER INFORMATION: encoding residues D183-1546 with flanking vector
US-10-159-151-4
```

```

Query Match      11.1%; Score 758.5; DB 15; Length 373;
Best Local Similarity 48.9%; Pred. No. 1.7e-45;
Matches 161; Conservative 58; Mismatches 103; Indels 7; Gaps 3;

Qy      4 KIRVVRRKRPPLSELEKKKDSIIIVKNNCTLYIDEPKYVDMTKYIERHEFIVDKVFD 63
      42 KIRVCVRKRPPLGREGVRRGEINIIIVEDKETLLVHEKKEAVDLQYIILGHVFFDEVFGE 101
Db      64 TVDNFVYENTIKPLIILIDYENGCVSCFAYGOTGSGKTYTLMGSOPYGOSDTPGIFQYA 123
      102 ACTNODVVMKTHPLIOHIF-NGSNATCFAYGOTGAGKTYMTGTH-----ENPGLYALA 155
Qy      124 AGDIFFLINIDYDNDKMGIFISFEYIYCGLYDLQKRKVAALENKEKEVVKDKILR 183
      156 AKDIFRQLEVSQPRKHLFWMISFEYIYCGLYDLNRRKRLPAREDSKHMVQIVGLQELQ 215
Db      184 VLTKEELILKMDIGVLLRKIGVNSONDESSRSHAILNIDKIDINKNTSLGKIAFIDLGS 243
      216 VDSVELLEVLILKSGKSRSTGATGVNADSSRSHAVIQIQKDSAKRT-FGRISFIDLGS 274
Qy      244 ERADTVVSQNKQOTQDGANINRSLALKECIRAMSDSKNHI PRDSELTQVLRDIFVGS 303
      275 ERAADARDSRQRTKMEGAEINOSLALKECIRALDQEHHTPPROSKLTQVLRDIFVGS 334
Db      304 KSIIMINISPTISCCQOTLNTLRYSSRVK 332
      335 KTCMIANISPSHVATEHTLNTLRYADRVK 363
```

RESULT 11

```

US-09-849-602-20
; Sequence 20, Application US/09849602
; Publication No. US20030165834A1
; GENERAL INFORMATION:
; APPLICANT: Scianlan, Matthew J.
; APPLICANT: Old, Lloyd J.
```


APPLICANT: Stockert, Elisabeth
APPLICANT: Chen, Yao-Tsung
TITLE OF INVENTION: Colon Cancer Antigen Panel
FILE REFERENCE: L0461/7105 (JRV)
CURRENT APPLICATION NUMBER: US/09/849,602
CURRENT FILING DATE: 2001-05-04
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn version 3.0
SEQ ID NO 20
LENGTH: 725
TYPE: PRT
ORGANISM: Homo sapiens
US-09-849-602-20

Query Match 11.0%; Score 750.5; DB 12; Length 725;
Best Local Similarity 41.9%; Pred. No. 1.5e-44;
Matches 173; Conservative 76; Mismatches 143; Indels 21; Gaps 10;

QY 4 KIRVVRRPLSELEKKKDDIITVKNCTLYIDEPRYVDMTKIERHEFIVDKVFD 63
DB 258 RIVCVARRPLNKKELAKKEIDVISIPKCLLVHEPKLVDTKYLENCAFCDFADE 317
QY 64 TVDNFTYENTIKPLIIDLYENGCVCGFAYGOTSGKTYMTLGS-QPYGSDTPGIFQY 122
DB 318 TASNEVYVYRFTARPLVQTFIEGQ-KATCFAYGOTSGKTHMGDLSGKAQNAKGIYAM 376
QY 123 AAGGIFFTPLN-IYDKNTKGFISFYIYCGKLYDILQKRWAALENGKEVYVVDLK 180
DB 377 ASRVVFLKNDPCYRKGLR-VYVTFEYINGKLFLLNKAKKRVLEDGKQOVVGLQ 435
QY 181 ILRLVTEELIKMID-GVLRKIGVNSQNDSSSHA1LINDLKINDKNTSLGKIAFD 239
DB 436 EHLVNSAD-VIKMLDKMSACRTSGQTFRANSSSHACFQILLR-AKGMHGFSLYVD 492
QY 240 LAGRGADTVSOKKQOTGDANINRSLLALKECIRAMSDKNHLPFDSLETYKLRDIF 299
DB 493 LAGNERGADTSARQTMGAEIKSLALKECIRALGQNAHPFRESKLTQVLRSDF 552
QY 300 VGR-SKSMINISPTISCEQTLNTLFYSSRVKFKKSCINEDDTNTERISIDSK 358
DB 553 IGENSRITMIAVISGISSCEYTLNTLRYADRVKELSPHSGPSGQ-----LIQME 603
QY 359 GSENNASSIENVVIXSNHLSNNNNKINRGKINDKINRNNILKKKSPDKPRE 411
DB 604 TEEMACG-NGALIPGLSKSEBELSSQMSFNEAMQINLEBKAMELKE 654

RESULT 12
US-09-723-276-2

Sequence 2, Application US/09723276
Patent No. US20020101952A1
GENERAL INFORMATION:
APPLICANT: Beraud, Christophe
APPLICANT: Freedman, Richard
TITLE OF INVENTION: No. US20020101952A1 motor proteins and methods for
FILE REFERENCE: 1032
CURRENT APPLICATION NUMBER: US/09/723,276
CURRENT FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 09/675,227
PRIORITY FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 319
TYPE: PRT
ORGANISM: Human
FEATURE:
NAME/KEY: VARIANT
LOCATION: (154)...(183)
OTHER INFORMATION: Xaa = any amino acid
US-09-723-276-2

Query Match 10.2%; Score 696.5; DB 10; Length 319;
Best Local Similarity 45.8%; Pred. No. 3.2e-41;
Matches 149; Conservative 51; Mismatches 118; Indels 7; Gaps 3;

QY 4 KIRVVRRPLSELEKKKDDIITVKNCTLYIDEPRYVDMTKIERHEFIVDKVFD 63
DB 1 KIRVCARRPLNKKELAKKEIDVISIPKCLLVHEPKLVDTKYLENCAFCDFADE 60
QY 64 TVDNFTYENTIKPLIIDLYENGCVCGFAYGOTSGKTYMTLGS-QPYGSDTPGIFQYA 123
DB 61 ACTNODVMKTTPELPHIF-NGGNATCFAYGOTSGAKTYMTGTH-----ENPGLYALA 114
QY 124 AGDIFFTPLNIDYDKNTKGFISFYIYCGKLYDILQKRWAALENGKEVYVVDLKILR 183
DB 115 AKDIFRQLEVSQPRKHLFWWISFYIYCGQLYDLNRRKXXXXXXXXXXXXXXXXX 174
QY 184 VLRTEELIKMIDGVLRKIGVNSQNDSSSHA1LINDLKINDKNTSLGKIAFDLAGS 243
DB 175 XXXXXXXXXVILKSSKRSSTGANGVADSSRSHAVIQIQLDSAKRT-FGRISFTDLGS 233
QY 244 ERGADTVSOKKQOTGDANINRSLLALKECIRAMSDKNHLPFDSLETYKLRDIFVYGS 303
DB 234 ERADNADSPRQTMGAEIKSLALKECIRALDQHTHTTFRQSKLTQVLKDSFIGNA 293
QY 304 KSIIMANISPTISCEQTLNTLRYSS 328
DB 294 KTCMIANISPSHVATEHTLNTLRYA 318

RESULT 13
US-09-820-843A-108

Sequence 108, Application US/09820843A
Publication No. US2003003963A1
GENERAL INFORMATION:
APPLICANT: Council of Scientific and Industrial Research
TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTE
FILE REFERENCE: 063915
CURRENT APPLICATION NUMBER: US/09/820,843A
CURRENT FILING DATE: 2001-03-30
NUMBER OF SEQ ID NOS: 118
SOFTWARE: PatentIn version 3.0
SEQ ID NO 108
LENGTH: 807
TYPE: PRT
ORGANISM: Plasmodium falciparum
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: hypothetical protein
NAME/KEY: misc_feature
OTHER INFORMATION: g1|3845292
US-09-820-843A-108

Query Match 10.2%; Score 696; DB 11; Length 807;
Best Local Similarity 26.6%; Pred. No. 1.2e-40;
Matches 253; Conservative 151; Mismatches 300; Indels 246; Gaps 43;

QY 328 SSRVKNFKNKSTCINEDDTNTERISIDSKGSENNASSIENVVIXSNHLSNNNNKIN 387
DB 10 SSRKNEEN-DIINKCDNS-----KINGKE-NIFAQKVGINSGHMSNDINK-N 58
QY 388 RGKINDKIERNNILKKKSPDKPREGFTSTFGKYSLSLNDIDKIKKKKKGLINYKSTLYN 446
DB 59 QEKKKKKKKKTHKYNINNTHTNDKNGG-QDINKPEYIERNNIINIKNDTNNI 117
QY 447 -DNTINKGNNNNNDNDNDNNNDNNNDSSSVYNN-----MINEMINNNI----- 495
DB 118 LDSSYNEEGNENRNDINNN 177
QY 496 -NNNINNN 554
DB 178 NNEINNTLLNNKNN 205

[illegible]

	Query Match	10.2#	Score 694.5	DB 11	Length 861	
	Best Local Similarity	26.4#	Pred. No. 1.6e-40			
	Matches 254	Conservative 145	Mismatches 325	Indels 237	Gaps 41	
QY	393	DKIERNNILKQKSFDEKREGEFTSTFGKYSLLNDIDKIKNKKGLINYKSTLYNDNTINK	452			
Db	21	DILNKYCFIPERKYNKFCGG-----NKY-----IPRBSNNNNNIIGNVNMGNMFVLLNN	69			
QY	453	KHNN	511			
Db	70	NNNNNRRLNTYNN	123			
QY	512	NNSHNNHLPOPNYAFTDTSFSDMSLDDMCHLNNDKSIPLFKKQJLRIJLKLRSSC---	568			
Db	124	NNNHFNIHNIIDYDSDYVKGGRHNGYVSSLSLNINNGYF---KXLDNN-----CYNL	172			
QY	569	--DNIMKKKKNNHLA-----RHSVGSFLTFMSYDPOKKONDFPKSNINMKED	615			
Db	173	PTNNLYIDKEGKMHLTGEKHYNAASNBRYNNNNKQTN--NYNNNSYNNNNFCCNNYNDNNY	231			
QY	616	NTPRD-----IIYESHVSNNN-----GNVLGLMKNTHDI	647			
Db	232	NNSNNKMGKNTYESLAIYLLKKEHMDVDEYNNKGNIRKDSSEKYNWDFPLYSKKANDI	291			
QY	648	ST-----KDEHNNDKILNNGVINI-----INNSVNSINNSNNSINNSNNSNSIYKSY	698			
Db	292	FTLADIKKYAKNKKKGGKNNKYMMNHNDNNSNS--NVLNNNNMS--NSN--NYNNIFKND	347			
QY	699	NSNSISDVGQRRYVNEMDTSKKKNDNIFPDALISCDNNM---YRITNNNNNNNN--NN	752			
Db	348	EEHTJTSN---FAKWFKNNNNNNNNNNNNTDITIKYLNKNSQSGHSDGKNNNNNNNNNN	403			
QY	753	NNNNNIIVENYNNNDGTNNSMKLYAAYNSHLFOPDNKTKTNIQNIPTNNKNOGNAVYS	812			
Db	404	NSNNKNNIFQGSNSYENWYNNIINNNNNNNNIISNNKKEASNTDNIITNSGREBEKIS-	461			
QY	813	MNFCHYLNLDKQVLIIDLNKEQKDKNIGHCDNNIIONPND---FEKKKGTNYFNNNNIV	868			
Db	462	-----NTVAELMKQ-----ISMIKERNKGLDYLEKKNTFGFLDNN---	497			
QY	869	IYNNNMGNNSPRAKYGLCSGHTSIDMKNNEMKNNEMKNNEMDNHKSNNNSSSSSS	928			
Db	498	--YQYSGSNV-----SSLEKNNKENDIYSKASAKIM-DIFRTLSNGLVQSES	545			
QY	929	-----SNNNIYNNINDDTFQNDYCHANDTFTRKKNNTINSNIYONDDIYTINS	980			
Db	546	LLVNGSVLNN	595			
QY	981	LNDYMSNTLHFKEKTYPTLSTVEDLYNKEMBEKHRLDDQKYDDNDNNNNVNNNNKNN	1040			
Db	596	NNYKNNKHKXSHMDNVYKKIFIN-----NYSNNDGNN--NSNNSN	634			
QY	1041	VDNNVD-----NN--NVNNNVND-----NDKNNVNNNNVNDDDVDVDFHNIKFPNNNEYS	1089			
Db	635	SNNNVHEYNNNNKNGFKNNKINNTYNNLEPDKNKMMNNNNNTYNNINKNNLSMNEFPFS--LS	692			
QY	1090	YFOKQVDTIINCLNSLIDISSMYDTEKILNINLISLKAYKAKQNVYKYYINEDIKMSLE	1149			
Db	693	F--NNSDINKNNKNOGINIT-----PIINSIL--RLDNEVDVHNNSSISENTQNAKVS	741			
QY	1150	EIDTASISYKRRVLLTLKLLLFKKVNDQINNETSDLKRLVMCHCNNPPDQGFHY	1209			
Db	742	NVLDLSKSLKASQGN-----NYNIPKRFN-----NNNNNNNSKFI	781			
QY	1210	AYSRLKDIINLIMLRQIWCESENNLRLLYOFLVVEYONKANSVILNVSNNGDIIILNK	1269			
Db	782	NYNS-----QQYPSHQOQOQOHO-----QQOQOQOQOQTLQIQTINSHLNDFNK	826			
QY	1270	K 1270				
Db	827	K 827				

Db	136	PRSEY-----LFSLDREKEKAGAGKSP/LCKCSFIEINQOYDLDBSAGLYIR	188
QY	168	ENGKEEYVYVKLIIRVLTKEELILKMDGVILRKIVGNSQNDSSRSHAILINDKODIN	227
Db	169	EHKKGVAVVAGVEOVVTSAAEAYOVLSGGEMRRRVAVSTSMRRESSRSHAVFTITESME	248
QY	228	KNTSLGKJA-----FIDLAGEBGADPYSONQOTQDGINIRSLATKECIRAM-----D	278
Db	249	KSNEIVNIRTSILNLVLDLAGSEROKOTHAEGRLKEAG-NIRKSLSCQVITVALVDGN	307
QY	279	SDKOHIPPRDSLETYVLDFIFGKSGSIMIANISPTISCCOEQTLWTLRRYSRVKPKNKS	338
Db	308	GKQIRAVCRDSTLFTLNRDSLGKNAKTALIAVHVGSRCFGFTLSTLNAQAQAKILIKKA	367
QY	339	TCINEEDDTNERI-----	358
Db	368	V-VNEDTOGVNSQ/LAEVYKLEQ/LAE/LASGQTPRES/FLTRDKKKTNYMEVFEQAM/LPFX	426
QY	359	GSEMASS-----IENVVTKSNHLSNNN-----N KING-----	389
Db	427	KSQEKSLIEKVTOLEDLTLLKEKFIQSNKMIYKFPREDQILREKLNHESGFLPEBO	486
QY	390	-----KINDXIE-----RNNILKXKS-----FDKPRGFGSTGK	419
Db	487	DRLSLAEINETO/LBQIEHHRVYAKYAMENHSLBENRRLYLBPVKRAQMDAQITAK	546
QY	420	---YSLINDIKIKKN-----KKGILNYKSTLYNDNTINKGN	455
Db	547	LEKAPSEISGKSPDKOGFSPKAKQEPCLPANTYK/LKAQ/LQ/OTEL-----	595
QY	456	NNNNNNNDNNNNNNNNNNNNDDSSVYANNINMHINNNINNNINNVN-----NNNNNNNN	513
Db	596	---NNSKOEYEEFKELTRKQLELSELOST--QKANNLEMLBAYACKRQEVSO/LNK	650
QY	514	SHNNHL-----PQRYUAFYDTSDFSLDDMCHLNNNNKSIPLHKONLD	558
Db	651	IHAETLKITTPYKAYOLHSRPVKLS-PEMGFSFSLYTQNSSILDN-----ILNE	701
QY	559	NIKLKNRSSCONNNKKNLHLARHVSGLKTMESYDPQKNKQDTPFKSINIKME-DNT	617
Db	702	PVPPEMNOAFPAISIELRTVOEQMSALQAKL-----DEBEHK-NLKLQOHVDKLEHHS	755
QY	618	PKDIIYBSRNV-SMMNGVVLGLM-----KXTHHDISTYDENHNDKINGVINII	667
Db	756	QWQELFSSERIDWTQOEBELLSQ/LNVLEKQLOETQKDNF-LKSEVHD-----LRVY	806
QY	668	NNSVYNSLNNSMNSINNSNNKNSIYKSNVYNSQ-----SISDQIYVNMEDTSNKN	722
Db	807	LHS-----ADKELSSVYLEYSSFTKQNEKEFKNLSEHHMVOLQ/LDWLRLEN	853
QY	723	DNIFEDALSCDNMYPAITNN	782
Db	854	EKL-LESACQIQDSVDN/QEITKPEIIDQLSRLQ/LFKKENETLKS/DLNNWELLAEKSR	912
QY	783	LFPQDNKNTSNIQNIINTNNKNNQDGNVYNSMFCYINLNDKNYLLDLNNKQKXDNHGC	842
Db	913	-----NNKLSLOFEEDKENSKEILIKVLEAV-----ROEKQETAKC	949
QY	843	DNNI-----IONRDPEKKKKKTNYNNNNIYVYNNNGNNSPPMK	883
Db	950	EQQMAKVQKLEESLATEKVISSLEKSRDSDK-----VADLM--NQIOELR	995
QY	884	YGLGSGHSHSDNNKNNKNNKNNEMKCN-----EMKCN-----	914
Db	996	TSVEKETITDIK-QELKQDINCKNTSALVDBREBSVLKQEOVILDLKETLRIRISE	1054
QY	915	-----HKSNNNNSSSSSSNNNNIYNNIND-----DTFQNDYCHNDTFTIR	958
Db	1055	DIERDMLCEDLAHATEQ/LNMLTEASKSGSLQ/LSQAEELTKKEALIGELQHKLN-----QX	1110
QY	959	KNNNNINNIYQNDIITYINSIANDYMN-----TLHFKC-----	994
Db	1111	KEEVOCKNEYN-----FKRQOLEHVMDSSAADPOSPKTPPFOHAKLETQBOEIED	1165

```

QY 995 -----KYVYPLTSTNEDVYNKEM-----EGHIRLDDDDKXDDDDNNNVDDNNK 1038
Db 1166 GRASKTSLSEHLYTKLNDREYVNAELIRKKEQLREMEYLRIESQOLIEKN----- 1215
QY 1039 NNVDNNVDDNNNVDDNNVDNNNDKNNVDNNNVDDDDVDFFHNIKNNNNNEYLSPYQKNVDTI 1098
Db 1216 WLGGQGLD--DIKQKENSDDNHDDNOQLKNKEBE-----SIK-----ERLA----- 1255
QY 1099 INNCLNSLIDSSMTVDITKEILANILLSKYAEKQNVYIKKYNIEDIKMSL-EEIDKT--- 1154
Db 1256 -----KSKIYEEML--KMAADLEEVSALYNNKMECLAMTDEVERTOTL 1297
QY 1155 AQSTYERKVLTKTLTLFPKKNVDTOINNENSDLRKDLWNC-----HICNNPDPOFH 1207
Db 1298 ESKAFQEKQLRSKLJEMREPERT--SQEMEMIRKO-VCLLAELNGKLVGHOVLHOKIQ 1354
QY 1208 FYAYSRLKEDIINLIMLROIWCESENRLLYOFLVWEYQNK 1249
Db 1355 YVV--RLKENVRLAE-----DEKJRAENVFLPEKKRSBS 1388

RESULT 24
US-10-342-224-62
; Sequence 62, Application US/10342224
; Publication No. US20030162294A1
; GENERAL INFORMATION:
; APPLICANT: Nathalie Verbugsen
; TITLE OF INVENTION: Genes Involved in Tolerance to Environmental Stress
; FILE REFERENCE: CNN-012US
; CURRENT APPLICATION NUMBER: US/10/342,224
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: US/09/762,154
; PRIOR FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: EP 98202634.6
; PRIOR FILING DATE: 1998-08-04
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 62
; LENGTH: 1269
; TYPE: PRF
; ORGANISM: Arabidopsis thaliana
US-10-342-224-62

Query Match 6.6%; Score 454.5; DB 12; Length 1269;
Best Local Similarity 33.3%; Pred.No.2:3e-23;
Matches 129; Conservative 80; Mismatches 137; Indels 41; Gaps 15;

QY 1 MNSKIKVYVVRKRPSELEKKKSDITTVKNNCTLYIDEPYKVDMTKYIERHEFIVDKV 60
Db 894 MKGIRIRYCHIRPLNEKESSEEREKQMLTVDEFTV---EIAWMD-----KRKHIIYDRV 945
QY 61 FDDTVDNFVYENTIKPLI---IDLVENGCVCSPAYGQTGSGKTYTMLGSPYQGSQDP 117
Db 946 FDMRASODDIFEDT-KYLQSAVDGY-NVCI---FAYGQTGSGKTFTI-----YGHESNP 995
QY 118 GIPFYAGDIFFTLINIDKXNTKGIF-ISFY--EIYCGKLYDL-----QKRKNVALLENG 170
Db 996 GLTPRAYKEFL--NLIKRSKSPFESLKAIVWELVODTVLDLLPEKSARLLKLEIKQDS 1052
QY 171 KKEVVVVDKILRLVTEELILMIDGVLRLKIGVNSQNDSSRSHAILNIDDKINKNT 230
Db 1053 KGMVFNENVTTIPISTIEELRMLIERGSERRHVSQTNMNESSRSHILISVLESIDLQT 1112
QY 231 ---SLGIATPDLAAGERGADTVSQNKQOTQDGANINRSLLALKECTIRAMDSKXNHI PFR 287
Db 1113 QSAARGKLSTVDLAGSERVKKSGSAGCOLK-EAQSIKNSISALGDVIGALSSGNOHPIYR 1171
QY 288 DSELTLYKLRLIPFGKXSIMIANISPTISCEQOTLNTLRSSRVKNPKXN-KSTPCINEED 346
Db 1172 NHHKLTMLMSDLSGNAKATTLMPFVNVSPAESMLDETIVNSLLYASRVRTTVDPDSKHISKEM 1231
QY 347 TINTERI-----SILDSKGSEMANSSIE 368

```


Db 175 PDKGVVQGLSFHQPASAEQLLELTGRNRRHQPTDANATSSRSHAIPOIFVQKQDRV 234
QY 226 --INKNTSLGKIATIDLAGSERGADTVSONKOTODGANINRSILALKECIRAM---DSD 280
Db 235 PGLTQAVQVAKMSLIDLAGSERASTHAKGERLR-EGANINRSILALINVALADAKGR 293
QY 281 KNHIFPRDSELTLYLRDIFVQSKSIMANISPTISCCEOTLNTL 325
Db 294 KTHVPRDSKLTRLKDSLGANCRTVMIAISPSLSLTEDYNTL 338

RESULT 29

US-09-847-874A-1
; Sequence 1, Application US/09847874A
; Patent No. US20020127668A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: KINESIN-LIKE MOTOR PROTEIN
; FILE REFERENCE: PR-0593 US
; CURRENT APPLICATION NUMBER: US/09/847.874A
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 09/467,946
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 1103
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 1281811
US-09-847-874A-1

Query Match 5.8%; Score 395.5; DB 10; Length 1103;

Best Local Similarity 31.3%; Pred. No. 2,7e-19;
Matches 126; Conservative 71; Mismatches 156; Indels 49; Gaps 16;

QY 3 SKIKVAVKRPPLSELEKKKSDITVKNKCTLYIDBRVYKVMTKYLE-----RHEFI 56
Db 4 ASVKAVAVRPFNARETSQDAKCVSMQNTTSILN-EKQSKDAKSFDPDYSVMSHT 62
QY 57 VDKVFDVTVDNFTYENTIKPLIIDLYENGCVCSFANGGCGSKTYYMLSGPQSGDT 116
Db 63 EDPGF---ASOOQVYRDIGEEMLHAFEGYNVC-IFAVGOTGAGKSYTMGRQEPGQO-- 116
QY 117 PGIPQYANGDIPTFLINIDKONTK---GIFISFYEIYQCKLYDL--OKRKVVALENGK 171
Db 117 -GIYPOLCEBDFS--RVSENOSAQLSYSEVSYMEIYCEYVADLLNPKSRGSLRVREHPI 173
QY 172 KEVVVKDLKILRVLTKEELILKMLIDGVLLRKIGVNSONDESSRSHAIINI-----DL 223
Db 174 LGPVYQDLSKLAIVSYADIDLMDCGNKARVAAINNMETSRSRAVFTIYVTOCHDOL 233
QY 224 KDINKNTSLGKIATIDLAGSERGADTVSONKOTODGANINRSILALKECIRP---MDS 280
Db 224 TGLD--SEVYSKSIIVDLASGER-ADSSGARGMGLGEGANINKSLTTLGKVISALADMSK 291
QY 281 K---NHIFPRDSELTKVLRDIFVQSKSIMANISPTISCCEOTLNTLRYSRVKNPKK 337
Db 292 KRKSDPTIPRDSVLTWLKENTIGNSRTMIALSPADINVEETLSTIRYADRTKOIR-C 350
QY 338 STCINEEDVTTERISILDSK-----GSENNASSIENV 370
Db 351 NAIINE--DPNARLIREQEEVAVARRELLMAGSLASALEGL 390

RESULT 30
US-10-167-831-2
; Sequence 2, Application US/10167831
; Publication No. US20030166209A1

; GENERAL INFORMATION:
; APPLICANT: Shuman, Stewart
; APPLICANT: Ho, C. Kiong
; TITLE OF INVENTION: Pharmacological Targeting of mRNA Cap Formation
; FILE REFERENCE: D6388D
; CURRENT APPLICATION NUMBER: US/10/167,831
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 09/752,165
; NUMBER OF SEQ ID NOS: 104
; SEQ ID NO 2
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: PEPTIDE
; OTHER INFORMATION: Plasmodium falciparum RNA triphosphatase Prt1
US-10-167-831-2

Query Match 5.7%; Score 388.5; DB 12; Length 596;

Best Local Similarity 21.5%; Pred. No. 3,7e-19;
Matches 171; Conservative 107; Mismatches 245; Indels 271; Gaps 33;

QY 372 IKSNNHLSNNNNNNKINCKINDKIERN---NIIKNSFDC-PREGFTSTFGKYSLLNDID 427
Db 24 LSGNIIILAFDHNHINKDIOIEGRVGLVIDKKNRIKLPTIDAIENNSDPO--A 81
QY 428 KIKKKKKGLINYKSTLYNDNTINKKNNNNNNNNNNNNNNNNNNNNNNNNNNNNSSSVNMI 487
Db 82 GIDRESEFELLDYFHNNTLKKRLSIRNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 130
QY 488 NHHNN 544
Db 131 -----NNNHHSNGNNTNQTSHYDADDKPTCNYSY----- 164
QY 545 NDKSIFLHKCN--LRDNIKLNSSCCDINMKKKNNLHLARHSGSLTWFSDPOKND 602
Db 165 -----DKGNACIYDFLELTKTSIDKYVIKNN-----SRIRTTY----- 201
QY 603 NTFPKSNINKMEDTPKOLYESHNVSNNGNVLGLKNTHTHDISTDENHNDKINNG 662
Db 202 ----- 201
QY 663 VINIINNSVNSINNSNNNSINNSNSIYSKSNVNSQISDVQIRYVENMDTSKN 722
Db 202 ---LNDNKGFTESMMIQLKDNLTIMVYTG-----NN 233
QY 723 DNIFPDAISCDNNMYPNITNN 782
Db 234 YDYFDDDEEDDDY-----NNNNNNNNNGDGTCTKTNAT--NTHGLTTSKSGHIYN--N 284
QY 783 LFQPDNNKNTSNICNITNKNKODGNVNSNF-CHYNLNDK---NYL---IDLNKE 833
Db 285 LV--DKDISIDYRISINIEYTKPIKSLYSKNTFVHERLKERFTFINTYVGLQVDMKIK 342
QY 834 OKDNHIGCD-----NNIQNRNDFEKKKKTFFNNNNNIVYVNNMGNNSPRKAYGLCG 888
Db 343 TKNNELVELEIPSKTIFKMSNLRNKKOSNYH-----FTGSLVNNR-----GIC- 391
QY 889 SHTSIDMKNNEMKNNEMKNNEMKNNIKNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 945
Db 392 -----SGLNVFKSKMLKNTM---ITLANN--SNGQNNLSLRLPHRPDDTISSEK 439
QY 946 ---DYCN-----DNTFTIRRNKNTNINSNIYQNDIYTTNSLNDYNSNTLHFEK 995
Db 440 EKFKYIHSVLPVIGDVYRVVTKGKIKRKI--KQDLTNKKEKINIKNVNDIRRNK 497
QY 996 YTPPLSLNE-DYKNKEGKH-----IRLDDQKXDDND-----NNNVNNNNKNVND 1043
Db 498 KSLQTI--NEVHVNKKMAFRGRTKIEVLLCSDEEYQNDVODINNEYDYQKNEEDT 555
QY 1044 NVDNNNV--DNNVNNNDKNNVNDNNVNDDDVDVFNHINKFNNEBYLSYFOKNVDTIINN 1101

Db 590 TSTKDALNGAKLAEAARAKONIGTLNHTNAQRTDLEGOINQATTVGVNTVKTANT 649
Qy IYKSNVSNOSIS--DVQIRYVENMDISNKNNDNIFPAISCDNNMNYININNNNNNN 749
Db 650 LDGMANSIQSGINDKDATLRNQNYLD--ADESKRAVYTOAVPAABGILNKKQIGNTSKADV 708
Qy 750 -----NNNNNNNNIDVENYNNRDTGNSMKLYAYN--SHNLFOPDN----- 788
Db 709 DNALNAVTRAKALNGADNLNNAKTSANTIDGLPNLTQLOKDNILKHOVEQAQVAVGVNG 768
Qy 789 ---NKATSN-----TONINTKNNO-----DGVNYSWNEFCHYN--LNDKNVLIID 828
Db 769 VKDKCNTLTAMGALRTSIQNDNTTKTSQNYLDASDSNKN-----NNTAVNNANAGVIN 822
Qy 829 LNKQKQXKN-----IHCQDN-----NIIQNRNDFEKKKKNF----- 861
Db 823 ATNPNMDANALINGMANOVNTTKALNGAONIAQAKTNATNTINAHDLNQAQKALKTQ 882
Qy 862 YNNNNIVLVNNMGNNSPRMKYGLCSHTSIDNNKNNEMKNEMKDNHKSNNN 921
Db 883 VNNAQVSDANNV--QHTATELNSAMTAKAIIADKERIKASGNVYVNAQEKQAVDSKYT 941
Qy 922 NSSSSSSNNNNIYNNINDDTFFQNDY-----CHNDTFTTRKNNNTINSNIYQNDII 975
Db 942 NAENIISGTPNATLTVDNVSAAQVNAKTAALNDGNNMLRVAKHEHANNITIDGLAQ----- 996
Qy 976 YTINSIDYMENTLHFEKXT-----YPLSTNEDIYKKEWKGKRIILDDQKXNDNN- 1030
Db 997 -----LNNAQAKLKEOVASATTLDGVOYTVKNSQTLTAMKGLDSTIANEATTIKAGONY 1051
Qy 1031 NNVDNNNNKQNV-----NNVDNNNNVDNN-----VDN 1056
Db 1052 TDASPNNNEVDSAVTAKAIIINOINSNFTMEPNTITQVTSQVTTKEQALNGARNLAQAKT 1111
Qy 1057 NDKNNVDN-----NNVDND-----DDVD-----FINIKK--FNNNERYL 1088
Db 1112 TAKNNLNNLTINNNAQKDALTRSIDGATTVAGVNOETAKATELNNAMHSLQNGINDETQT 1171
Qy 1089 SYPOKNVNTIINNCLNSIDISMYDTEILNINILSYKAEKQNVIKKYNEDIKHMSL 1148
Db 1172 KQTKRYLDAEPSK-----KSAYQAVNAKAIILT--KASGVNDVRAVQEAQALQVNA- 1220
Qy 1149 EIIDTKAQSIYERKVLTKLLFKRVNDTQINNSTDRLKDL--VMCHICN 1199
Db 1221 -----STKTAL-----NGDAKLNKAARAKQTLGTLRHINN 1251

RESULT 36
US-09-925-300-1228
; Sequence 1228, Application US/09925300
; Patent No. US20070151681A1
; GENERAL INFORMATION:
; APPLICANT: Steve Rosen,
; APPLICANT: Craig Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1228
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (75)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE

; LOCATION: (147)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (435)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1228

Query Match 4.9%; Score 336.5; DB 10; Length 460;
Best Local Similarity 28.7%; Pred. No. 1,2e-15;
Matches 100; Conservative 71; Mismatches 145; Indels 33; Gaps 10;

Qy 3 SKIRVVRRKPLSELEKKKDSIIIVKNNCTLYIDEPKYVDMTKYTEHHEFVQVFD 62
Db 20 ARVAVRALRPFDVGTAGADPPCVRGMDSCSLSIAMWRHNOETLKY-----QF--DAFYG 73
Qy 63 DTYNFVYENTIKELIIDLNYENCVCSCRAYGOTSGKTYTMGSGQPYGSDPRGIFQY 122
Db 74 EXSTQDDIVAGSVOPILRHILE--GQNASVLAYGEGTGKHTMLGS-----PEQGVYIPR 127
Qy 123 AAGDIFELNIYDKNTKG-----IFISFEIYCGLYDLQKPK--MVAALENGKEEV 175
Db 128 ALMDL---LQUTREGAEGRPWKLSTVMSYLEIQEKVLDLDPASGDVIREDCRNIL 184
Qy 176 VKDKILRVLTKEELIKMIDGVILRKIGVNSQNDSSRSHAIINIDKINKTSL--- 232
Db 185 IPLGSKPISFADFERRHFLPASRRRTVGATRLQSRSSRAVLD--LVKVDQRERLAPF 241
Qy 233 ----GKIAFLDLASERGAQTVSONKQOTQDGNINNSLLALKECIRAMSDKNHIFPD 288
Db 242 RQREGKYLIDLASSENRRNGKRLKESGA--INSLFVYGVVALNGLPFRVYRD 300
Qy 289 SELTKVLADIFVGSKSIMIANISPTISCEQTLNTLYSRVKNFRK 337
Db 301 SKLTRLQDSLGSAHSILIANINPERRFYLDYTSALNPAARSKEVINR 349

RESULT 37
US-10-171-311-196
; Sequence 196, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumel
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatic, Karen
; APPLICANT: Gannavara, Manjula
; APPLICANT: Hoerish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 196
; LENGTH: 890
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-196

Query Match 4.9%; Score 333.5; DB 15; Length 890;
Best Local Similarity 25.7%; Pred. No. 4.7e-15;
Matches 117; Conservative 64; Mismatches 128; Indels 147; Gaps 15;

THIS PAGE BLANK (RSP10)

XX 07-MAR-1991.
PD
XX
XX 17-AUG-1990; 90WO-US04651.
PR
XX 18-AUG-1989; 89US-0395732.
PR 26-FEB-1990; 90US-0485158.
PR 18-JUN-1990; 90US-0540363.
XX
PA (SCRI) SCRIIPS CLINIC & RES FOUND.
XX
PI Dyer CA, Curtiss LK, Smith R;
XX
DR WPI; 1991-087246/12.
XX
XX Polypeptide(s) copying apo-1lipoprotein E induced cell
PT differentiation - useful for treating hypercholesterolemia,
PI inhibiting ovarian androgen secretion and for prodn. of diagnostic
PT systems.
XX
XX
PS Claim 2; Page 98, 123pp; English.
CC
CC A polypeptide capable of inhibiting lymphocyte proliferation and/or
CC ovarian androgen secretion comprises a plurality of segments, as
CC indicated in the features. The segment may be repeated 2-10 times.
CC In this example the segment is repeated 10 times. The N-terminal
CC may comprise NH2 or a leader segment of 1-20 amino acids, the
CC C-terminal comprises COOH or a tail segment of 1-20 amino acids.
CC Between each segment is a first and second spacing segment of
CC 1-20 amino acids. P(141-155) is a peptide obtained from apo E.
CC (See also ARI1107 and ARI1109.
CC (Updated on 25-MAR-2003 to correct PA field.)

Query Match	13.7%;	Score 940;	DB 12;	Length 590;
Best Local Similarity	32.4%;	Pred. No. 4,4e-4s;		
Matches 235; Conservative	99;	Mismatches 215;	Indels 176;	Gaps 18

```
Query Match          13.7%; Score 940; DB 12; Length 590;  
Best Local Similarity 32.4%; Pred. No. 4,4e-45;  
Matches 235; Conservative 99; Mismatches 215; Indels 176; Gaps 18
```

QY 374 SNHLLSNNNNKKINRGIINDKIERNNILNKYSFDPKREGFTSTFGKSYSLNDIKIKNK 433
 :|::||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::
DB 9 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNLRLKRKL-----LRDADLNNNN 59

QY 434 KKGGLINYKSTLYNDITIRKGNANNNNNDDNDNDNNNNNNNDSSM----- 482
 ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 60 -----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNRKLRLRDA 107

QY 483 --VNMIMNMIMNNIIINNIVANNNNNNNNNNNNSHNNLPORPIYAFTPDSLSLDMMC 540
 ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 108 DDLNN----- 145

QY 541 HLANNDKSFLFKAILRDNIKLKNSSCDNIRAKKKNNULHLARHSVGSKLTMFSTYPCK 600
 ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 146 -NNNNNNLRKLRLRADDLNNNNNNNNNNNNNNNNNN-----N 182

QY 601 KDNTPEKSNIINKMEDTPTPOLYESRYNSWNGNVTLGINKUTHHDISTKDEHNDRKIN 660
 :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 183 NNN---NNNNNNNN--NNNNNNNNNNLNRLRKRLLRADDLNNNNNNNNNN 230

QY 661 NGVINIIINSYNVSINNSWNISINNSMNWSIYSKYNSNOSISDVQIRYVENEDTSNK 720
 |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 231 N-----NN 277

QY 721 NNNDITPFDAISCDDNYTPYITNNNNNNNNNNNNNNNNNNIDVERYYNPDGTNSKLYAYS 780
 ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 278 NNNN-----NN 324

QY 781 HNLFPQDNKNTSNIIONTNNKNODGVNYSWNPCHAYLANDKTYLI-----D 828
 :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 325 RQADDLNNNLRKLRLRDDAD 384

QY 829 LANKEOXKHIGCDNNIIOQRNDFEKKKTTFYNNNNVIYVNMMGNNSPMRTKGLCG 888
 ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:

[illegible]

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139219.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.

[illegible]

Db 652 -----REKCDPRQNMGRKRVSS-----SD 671
QY 499 INVNNNNNNNNNNNNNHNPQPNVAFPTDTSDFSSLDNMCHLNNDKSTFLHKRLRD 558
Db 672 IPTLTNRQVASEFASROYETASROYE-TDPSLDENIDAL---LEEEFALIAHREKIED 727
QY 559 NIKL 562
Db 728 TWEI 731
RESULT 4
AAAG41921
ID AAAG41921 standard; Protein: 814 AA.
XX
AC AAAG41921;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 52217.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.


```

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149426.
PR 18-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151086.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 11.5%; Score 786; DB 21; Length 814;

Best Local Similarity 34.8%; Pred. No. 3e-36; Mismatches 103; Indels 108; Gaps 21;

```

QY 3 SKIKVVRKRPPISELEKKKKSDIITVKNKNTLYIDEPYKVDMTYIERHEEIVDKFVD 62
Dp 212 AKIKVVRKRPPISELEKKKKSDIITVKNKNTLYIDEPYKVDMTYIERHEEIVDKFVD 270
QY 63 DTVDNFTYENTIKPILIDLYENGCVSCGAYQOTSSGKTYTTLGSGOPYQOSTPGIFQY 122
Dp 271 EDVSNDEVYRATIEPIIPILIFOR-TRATCFAYQOTSSGKTYTTLGSGOPYQOSTPGIFQY 318
QY 123 AAGDIFLPLN--IYDKNTKGIPISEFYEITYCGLYLLOKRAKVAALENKKEVVVDLK 180
Dp 319 AVEDLMKRLRQPYVYSQNRK-LMLSTFEITYGTLFLLSRKTLCKREDRQOVCYVLQ 377
QY 181 ILRLTKRELILKMD-GVLLRKIGVNSQNDSSRSRSHAIL-----NIDLKQIKOTS-- 221
Dp 378 EYEV-SDVQIVKDFIKGNAERSTGTGANBESSRHAILQLVKKGHVEKQDTRRRNDS 436
QY 232 -----LGIKAFIDLAGSERGADTVSOKKOTODGAMINRSLALKECIRANDSDKNI 284
Dp 437 NELPGKRVKGISFIDLAGSERGADTVSOKKOTODGAMINRSLALKECIRANDSDKNI 496
QY 285 PPRDSELTQVLRDIFVYKSKSIIMIANISPTISCCBOTLNTLRYSRVK-----NFKNKS 338
Dp 497 PFRGSKLTVLRDSFVGNRSRTWISCTSPAGSCERTLNTLRADRVKSLSGNSGSKQD 556
QY 339 TC-----INEEDTNTERTISLDSKSEMNASSIENVVJKNHLSNNNNKINRQKIND 393
Dp 557 TANSMPVAKDPLLGNDVEDVEPEPQEVNPEPTRRVEKD---SNSSTSGIDFQPTV 613
QY 394 KIERNNILKQKSKDRREGFTSF-GKYSGLNDI-----DKIKK--NKKKGLI 438
Dp 614 YRESGI-PSFMDKRSSEPNSSFAGSTQQRNNISSYPOETSDREKVKVSPPRGKL- 671
QY 439 NYKSTLYNDNTINKKNNNNNNNDNDNNNNNNNNNDSSSMVNMIMHMINNNNN 498
Dp 672 -----REEKPRQWKSQDVSS-----SD 691
QY 499 INVNNNNNNNNNNNNNNHNLPOPNYAFTDTSDFSLDDMCHLNNDKSIPLAKKRLD 558
Dp 692 IPTLTFRQNASBTASROYETASROYE-TDPSIDENLDAL---LEBEALIAHREKIED 747
QY 559 NIKL 562
Dp 748 TWEI 751

```

RESULT 5
 ABB58104
 ID ABB58104 standard; Protein; 803 AA.
 AC ABB58104;
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 1104.
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 OS Drosophila melanogaster.
 OS
 PN WO200171042-A2.
 PD 27-SEP-2001.
 PF 23-MAR-2001; 2001WO-US09231.
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 PA (PEKE) PE CORP NY.
 PI Venter JC, Adams M, Li PMD, Myers EW;
 XX

DR WPI: 2001-6556860/75.
 DR N-PSDB; ABL02207.
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 1104; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutic and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB101875) and the encoded proteins
 CC (AB057737-AB0578022).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 803 AA;
 Query Match 11.4%; Score 783; DB 22; Length 803;
 Best Local Similarity 38.8%; Pred. No. 4,4e-36;
 Matches 193; Conservative 78; Mismatches 153; Indels 74; Gaps 13;
 Oy 4 KIKVVRKRPISLEKKKSDITTVKNNCTLYIDEPYKDKMTYIRHEFIYDKFPDD 63
 Db 276 QITVCVKRPIRSKREVRKEIDIVSPKDWLIVHEPRSKVDLTKEFNHFRDYAND 335
 Oy 64 TVNDFYVENTIKELIIDLYENGCCGCEAYGOTSGKTYMLISOPVG--QSPDPGIFQ 121
 Db 336 TCDNAMYKTKAKPLVKTIFEGG-WATFAVGOTSGKTHM-GGEFGKQDCKNGIYA 393
 Oy 122 YAAADIFTELNI--YDKNTKGIPISFYEYICGLYDLQKRWAALENGKEVAVMDL 179
 Db 394 MAADVFYTLMPRYRANLV-VSASFPEISGKVFDLISQKRLVLEDKQOVVGL 452
 Oy 180 KILAVLKEELIKMID-GVLLRKIGVNSQNDSSRSHAIINIDKIDINKTSIGKIAFI 238
 Db 453 -TERKVGVEEVLKLIQGNARITSGQTSANSNSRSHAVFOIVLRPGSTRIHGKFSFI 511
 Oy 239 DIASSEGADIVSONKOTOTGANINRBLAKECIRAMDQKHHIPRDELTQVARDI 298
 Db 512 DLAGNERVDTSADROTREGAEINSLALKECTIRLGGQSAHLPRVSLQVWLDS 571
 Oy 299 FVG-KSKSIMIANISPTISCEQTLNLTLYSSRVNPFNKSTCINEEDTTERISILDS 357
 Db 572 FIGKSKTCMTAMISPGLSCEHTLNTLRADRVKELVVKD-IVEVCPGSDTEPIEITDD 630
 Oy 358 KGEEMASSIENVVYKSHLSNNNNKINRGKINDKIERNNILKNKSFDPREBGFSTF 417
 Db 631 EEEEBE-----LNMVPHSHQHPNSHAPASQ----- 656
 Oy 418 GKYSSLNDIKTKKKKKGLINYKSTLYNDNTINKKNNNNNNNDNNNNNNNNNN 477
 Db 657 -----SNQORAPASHHSGAVT-----HNNNNNNKNGAHGMD----- 689
 Oy 478 DSSSVNNNMNNNNNNI 495
 Db 690 --LAWLSLSLSEHMSDEL 705
 RESULT 6
 AAB18250
 ID AAB18250 standard; Protein; 1844 AA.
 AC AAB18250;
 XX
 XX 07-NOV-2000 (first entry)
 DT
 XX Plasmodium falciptarum chromosome 2 related protein SEQ ID NO:107.
 DE

XX
 KW Plasmodium falciptarum; chromosome 2; human malaria parasite; vaccine;
 KW antimalarial; malaria; protozoacide; infection; insecticide.
 XX
 OS Plasmodium falciptarum.
 XX
 PN WO200025728-A2.
 XX
 PD 11-MAY-2000.
 XX
 PF 05-NOV-1999; 99WO-US26796.
 XX
 PR 05-NOV-1998; 98US-0107311.
 XX
 PA (HOFF/) HOFFMAN S.
 PA (CARU/) CARUCCI D.
 PA (GARD/) GARDNER M.
 PA (VENT/) VENTER J C.
 XX
 PI Hoffman S, Carucci D, Gardner M, Venter JC;
 XX
 DR WPI: 2000-365347/31.
 XX
 PT Proteins encoded by chromosome 2 of the human malarial parasite,
 PT Plasmodium falciptarum, useful as antimalarial vaccines and in the
 PT diagnosis of P.falciptarum infection -
 XX
 PS Disclosure; Page 246-250; 577pp; English.
 CC
 CC The present invention describes proteins and their fragments (I) encoded
 CC by chromosome 2 of the human malarial parasite, Plasmodium falciptarum.
 CC Also described are: (I) nucleotide sequences (II) encoding (I); and (2)
 CC vaccines against P. falciptarum infection comprising (I) or (II).
 CC (I) and (II) are useful for the development of vaccines against
 CC P. falciptarum infection. (I) and polyclonal antisera or a monoclonal
 CC antibody raised to immunogens comprising the sequences of (I), are
 CC useful in the detection of infection with P. falciptarum. Furthermore,
 CC (I) (especially when they are rifins or secreted or membrane proteins)
 CC can aid the identification of drugs to treat or prevent P. falciptarum
 CC infection, or they can be used to identify drug resistance in
 CC P. falciptarum. Sequencing of the Plasmodium chromosome 2 and the
 CC subsequent identification of proteins encoded by it will help to expand
 CC our understanding of parasite biology, a process hampered by the
 CC complexity of the parasite lifecycle, and provide new targets for
 CC vaccine and drug development. Parasite resistance to drugs and mosquito
 CC resistance to insecticides have led to a resurgence of malaria in many
 CC parts of the world, and there is a pressing need for vaccines and new
 CC drugs. AAT70078 to AAT70287 and AAB18144 to AAB18352 represent nucleotide
 CC and protein sequences given in the present invention, but which are not
 CC specifically mentioned within the specification.
 XX
 SQ Sequence 1844 AA;
 Query Match 11.4%; Score 778; DB 21; Length 1844;
 Best Local Similarity 24.5%; Pred. No. 2.2e-35;
 Matches 340; Conservative 202; Mismatches 439; Indels 406; Gaps 66;
 Oy 205 VNSQNDSSRSHALINI-DLKIDINKTSIGKIAFIIDLAGSRGADYVSQNKQOT----- 258
 Db 37 VNRNDNNYEMWYNSINGHMSNINNNTNODATINELHMYNSDKINIAHNNQVHMTA 96
 Oy 259 -----DGANINRSLLALEKCIKRAMSDKNHIFPRDSELTQVARDI FVGKSKSIMIANISP 313
 Db 97 TYNNMDKNNANNNIHIANHNH-NWTDQNT--FVNSTTNNGM-----NNTLKNNPNN 146
 Oy 314 TISCEQTLNLTLYSSRVNPFNKSTCINEED-TTERISILDSKGEEMASSIEN-- 369
 Db 147 NNN-----TVNNSFYHNTDNNELNFTRNQNEODTYVNNNIINNENNQVDDKINNINNPN 202
 Oy 370 -----VVIKSHLSLNNNNN-----KINGKINDKIERNNILKNKSFDPREBGFSTFGKYS 421
 Db 203 KNVESINKFNHITVMQNFNHFIPNISGK-NGNLNNASLSH-----NNVSVS 249

Db 72 ACTNODVYMKTHPLDQHIF-NGGNATCFAYGQTGAGKTYMTIGTH-----ENPGLYALA 125

Qy 124 AGDIFFTLNIYDKNTKGIFISFEIYCGLYDLQKRWAALENGKEVVDLKLRL 183

Db 126 AKDIFRLEVSQPRKHLFWMISFEIYCGLYDLNRRKRLFAEDSHHWQVIGLOELQ 185

Qy 184 VLTKSEELILKMDVLLRKIGVNSQNDSSSRSHAILNDLKDINKNTSLGKIAFIDLAGS 243

Db 186 VDSVELLEVLILKSKSKERSTGATGVNADSSRSHAVIQIKDSAKRT-FGRISFIDLAGS 244

Qy 244 ERGADTVSQRKQOTOGANINRSLALKECIRAMDSDKNHIPPDSSETKYLDIFVYGS 303

Db 245 ERADADSDRQRTMEGAELNQSLLALKECIRALDQETHTPFRQSKLTQVLKDSFIGNA 304

Qy 304 KSIIMANISPTISCCBOTLNTLRYSRVKFNKSTCINEEDPTN 348

Db 305 KTCMIANISPSHVATEHTLNTLRADRVKELKGIKCTSVTSRN 349

RESULT 8

AAU77184 standard; Protein, 381 AA.

AAU77184;

02-JUL-2002 (first entry)

Human Kint-3 DNA fragment with flanking vector sequences #2.

Human Kint-3 motor protein; Kint-3; microtubule; hyperplasia; depolymerisation; cellular proliferation disorder; cancer; restenosis; vascular malfunction; abnormal wound healing; cardiac hypertrophy; gout; inflammatory disorder; immune disorder; rheumatoid arthritis; psoriasis; diabetic retinopathy; neurological disorder; immunosuppressant; vasotropic; vesicular transport disorder; immunosuppressive; antiinflammatory; cyostatic.

Homo sapiens.

Synthetic.

WO200226929-A2.

04-APR-2002.

28-SEP-2001; 2001WO-US30750.

29-SEP-2000; 2000US-0675227.

(CYTO-) CYTOKINETICS INC.

Beraud C, Guo J, Freedman R, Patel UA, Davies KA;

WPI; 2002-352124/38.

N-PSDB; ABK47999.

An novel isolated microtubule motor protein, Kint-3 used for preparing compositions for the diagnosis, prevention or treatment of diseases such as cancers, hyperplasias or restenosis -

Disclosure; Fig 4; 68pp; English.

The invention relates to a human kinesin motor protein (Kint-3), with microtubule depolymerisation activity, and the polynucleotide encoding it. The novel Kint-3 motor protein can be used to prepare compositions for the diagnosis, prevention or treatment of cellular proliferation disorders, cancers, hyperplasias, restenosis, vascular malfunctions, abnormal wound healing, cardiac hypertrophy, inflammatory and immune disorders (such as rheumatoid arthritis, gout, psoriasis and diabetic retinopathy) neurological disorders and disorders of vesicular transport. This sequence represents a human Kint-3 polypeptide fragment with flanking vector sequences.

Sequence 381 AA;

Query Match 11.2%; Score 768.5; DB 23; Length 381;

Best Local Similarity 47.5%; Pred. No. 1,3e-35;

Matches 164; Conservative 58; Mismatches 116; Indels 7; Gaps 3;

Qy 4 KIRVYVARKRLPSELEKKKQSDIITYVKNCTLYIDEPRYVDMTKYTERHEPVDKYFDD 63

Db 30 KIRVCVARKRLPGRREVRGRGIIITVEDKETLLVHEKKEAVDLYTLQHVFFVEVGE 89

Qy 64 TVDNFYVENTIKPLIIDLYENGCVCSCFAYGQTGAGKTYMTIGTH-----ENPGLYALA 143

Db 90 ACTNODVYMKTHPLDQHIF-NGGNATCFAYGQTGAGKTYMTIGTH-----ENPGLYALA 143

Qy 124 AGDIFFTLNIYDKNTKGIFISFEIYCGLYDLQKRWAALENGKEVVDLKLRL 183

Db 144 AKDIFRLEVSQPRKHLFWMISFEIYCGLYDLNRRKRLFAEDSHHWQVIGLOELQ 203

Qy 184 VLTKSEELILKMDVLLRKIGVNSQNDSSSRSHAILNDLKDINKNTSLGKIAFIDLAGS 243

Db 204 VDSVELLEVLILKSKSKERSTGATGVNADSSRSHAVIQIKDSAKRT-FGRISFIDLAGS 262

Qy 244 ERGADTVSQRKQOTOGANINRSLALKECIRAMDSDKNHIPPDSSETKYLDIFVYGS 303

Db 263 ERADADSDRQRTMEGAELNQSLLALKECIRALDQETHTPFRQSKLTQVLKDSFIGNA 322

Qy 304 KSIIMANISPTISCCBOTLNTLRYSRVKFNKSTCINEEDPTN 348

Db 323 KTCMIANISPSHVATEHTLNTLRADRVKELKGIKCTSVTSRN 367

RESULT 9

AAU77182 standard; Protein, 1368 AA.

AAU77182;

02-JUL-2002 (first entry)

Human kinesin motor protein Kint-3.

Human, kinesin motor protein; Kint-3; microtubule; hyperplasia; depolymerisation; cellular proliferation disorder; cancer; restenosis; vascular malfunction; abnormal wound healing; cardiac hypertrophy; gout; inflammatory disorder; immune disorder; rheumatoid arthritis; psoriasis; diabetic retinopathy; neurological disorder; immunosuppressant; vasotropic; vesicular transport disorder; immunosuppressive; antiinflammatory; cyostatic.

Homo sapiens.

Key Location/Qualifiers

Domain 224..545

/label= Motor_domain

WO200226929-A2.

04-APR-2002.

28-SEP-2001; 2001WO-US30750.

29-SEP-2000; 2000US-0675227.

(CYTO-) CYTOKINETICS INC.

Beraud C, Guo J, Freedman R, Patel UA, Davies KA;

WPI; 2002-352124/38.

N-PSDB; ABK47999.

An novel isolated microtubule motor protein, Kint-3 used for preparing compositions for the diagnosis, prevention or treatment of diseases such as cancers, hyperplasias or restenosis -

OS Synthetic.
XX WO200226929-A2.
XX 04-APR-2002.
XX 28-SEP-2001; 2001WO-US30750.
XX 29-SEP-2000; 2000US-0675227.
XX (CYTO-) CYTOKINETICS INC.
XX Beraud C, Guo J, Freedman R, Patel UA, Davies KA;
XX WPI: 2002-352124/38.
XX N-PSDB; ABK47998.
XX
XX An novel isolated microtubule motor protein, Kln-3 used for preparing
PT compositions for the diagnosis, prevention or treatment of diseases
PT such as cancers, hyperplasias or restenosis -
XX
XX Disclosure; Fig 3; 68pp; English.
XX
XX The invention relates to a human kinesin motor protein (Kln-3), with
CC microtubule depolymerisation activity, and the polynucleotide encoding
CC it. The novel Kln-3 motor protein can be used to prepare compositions
CC for the diagnosis, prevention or treatment of cellular proliferation
CC disorders, cancers, hyperplasias, restenosis, vascular malfunctions,
CC abnormal wound healing, cardiac hypertrophy, inflammatory and immune
CC disorders (such as rheumatoid arthritis, gout, psoriasis and diabetic
CC retinopathy) neurological disorders and disorders of vesicular transport.
CC This sequence represents a human Kln-3 polypeptide fragment with
CC flanking vector sequences.
XX
XX Sequence 373 AA;
SQ
Query Match 11.1%; Score 758.5; DB 23; Length 373;
Best Local Similarity 48.9%; Pred. No. 4.5e-35;
Matches 161; Conservative 58; Mismatches 103; Indels 7; Gaps 3;
QY 4 KIKVVRKRPSELEKKKSDIITVKNCTLYIDEPYKVDWTKYIERHEFIYDKVFD 63
DB 42 KIRCVRRKRPGLMREVRGEINILITVEDKETLVHEKKEAVDLQYILQHFYFDEVEGE 101
QY 64 TVDNFTYVENTIKPLIIDLYENGCVCSGFAVGOTSGSKTYMTLGS-QPYGSDTPGIFQY 123
DB 102 ACTQODVYMKTHPLIQHIF-NGGNATCFAYGQTGAGKTYMTIGTH-----ENPELVALA 155
QY 124 AGDIFFTFLNIYDKNTKGIFFISFYEYCGKLYDLLQKRMVALENGKEVAVVKDLKILR 183
DB 156 AKDIFRQLEVSQPRKHLFWWISFYIYCGQLYDLNRRKRLPARBDSKHMVQIVGLQELQ 215
QY 184 VLTKKEELLKMIIDGVLLRKIGVNSONDESSSHALINIDLDKINKNTSLGKIAFIIDLGS 243
DB 216 VDSVELLEVLILKSGEKSTGATGVNADSSSHAVIQIOLKDSAKRT-FGRISFIDLGS 274
QY 244 ERGADTVSONKOTQDGNANINRSILALKECIRAMSDSKNHPFRSELTLYLRDIFVYKS 303
DB 275 ERADARDSRQTKMEGAEINOSLALKECIRALDOETHHPFRQSKLTQVLKDSFIGNA 334
QY 304 KSIMIANISPTISCEOTINTLRYSSRYK 332
DB 335 KTCMIANISPSHVATEHTLNTLRKYADRYK 363
Db
RESULT 12
ID AAY49949
XX AAY49949 standard; Protein; 730 AA.
AC AAY49949;
XX
DT 02-FEB-2000 (first entry)
XX

DE Xenopus laevis kinesin central motor 1.
XX
XX Strongylocentrotus purpuratus; katanin; p60 subunit; p80 subunit;
KW microtubule severing protein; ATPase; detection; identification;
KW microtubule depolymerisation inhibitor; screening; cell cycle;
KW kinesin central motor 1; XKCM1; regulation; antimitotic.
XX
XX Xenopus laevis.
OS
XX WO9953295-A1.
PN
XX 21-OCT-1999.
PD
XX 13-APR-1999; 99WO-US08086.
PF
XX 14-APR-1998; 98US-0081734.
PR
XX (REGC) UNIV CALIFORNIA.
PA
XX Vale RD, Hartman JJ;
PI
XX WPI: 2000-013108/01.
DR
XX N-PSDB; AA235759.
DR
XX
XX Identifying agents that modulate microtubule depolymerisation -
PT
XX Disclosure; Page 66; 69pp; English.
PS
XX The present invention describes a method for identifying a modulator of
CC microtubule (MT) depolymerisation by contacting a polymerised MT with
CC an MT severing or depolymerising protein and a candidate agent in the
CC presence of ATP or GTP, and detecting formation of tubulin monomers,
CC dimers or oligomers. The method is useful to detect agents that
CC regulate the cytoskeleton and the cell cycle. Such agents have potential
CC for use as anti-mitotic agents. The present sequence represents the
CC Xenopus laevis kinesin central motor 1 (XKCM1), which regulates
CC microtubule dynamics during mitotic spindle assembly. XKCM1 localises
CC to centrosomes and appears to regulate the polymerisation dynamics of
CC microtubules, and so is suitable for use in the method of the invention.
XX
XX Sequence 730 AA;
SQ
Query Match 11.1%; Score 758.5; DB 21; Length 730;
Best Local Similarity 42.1%; Pred. No. 9.6e-35;
Matches 175; Conservative 84; Mismatches 130; Indels 27; Gaps 12;
QY 4 KIKVVRKRPSELEKKKSDIITVKNCTLYIDEPYKVDWTKYIERHEFIYDKVFD 63
DB 262 RICEVVRKRPPLNKOELSKKEIDIVSPSKNIVLHEKRLKVDLTKYLENOAFRDFSPDE 321
QY 64 TVDNFTYVENTIKPLIIDLYENGCVCSGFAVGOTSGSKTYMTLGS-QPYGSDTPGIFQY 122
DB 322 TATNEVVYRTARPLVOSIFEQG-KATCFAYGOTSGSKTYMTMGDPFGSKQNVSKGYVAF 380
QY 123 AAGDIFFTFLNIYDKNTK---GIFISFYEYCGKLYDLLQKRMVALENGKEVAVVKD 178
DB 381 ASRVVFLLL--DQPRKXHDLDVFVTFEYFNKVPDLNKKTKYRLVEDAKQEOVYVG 437
QY 179 LKILRVLTKEELLKMIID-GVLLRKIGVNSONDESSSHALINIDLDKINKNTSL-GKIA 236
DB 438 LLEKQVLSADN-VEKMIIEISACTSGQTFANNTSSSHACLOILR--RGSLLHGFSS 493
QY 237 FIDLAGSERGADTVSONKOTQDGNANINRSILALKECIRAMSDSKNHPFRSELTLYLR 296
DB 494 LVLDLAGNERGVDTASADITMKGAELNRSILALKECIRALGQNKSHTPPRESLTQILR 553
QY 297 DIFVKG-SKSIIMIANISPTISCEOTINTLRYSSRVNFKXKSCINEEDPTNTERISIL 355
DB 554 DSFIGENSRCTCMIMLSGFGNSCEYTLNTLRKYADRVLELSPQNA-----ETNDNLQME 607
QY 356 DSKGSENNAGSIEENVVYKSNHLSNNNNNNKINRCKINDKIERNNIILNKKSDPKRE 411
DB 608 DSGGSH--ASIEGLQLODDPFLKDEBELSTHN--SFODALNRVGELEBDKAVDELRE 658


```

RESULT 13
AAE14504
ID AAE14504 standard; Protein; 405 AA.
XX
XX AAE14504;
AC
XX
DT 09-APR-2002 (first entry)
XX
XX Human mitotic centromere-associated kinesin protein fragment #6.
DE
XX
XX Human; MCAK; mitotic centromere-associated kinesin; motor domain;
KW cellular proliferation disorder; cancer; hyperplasia; restenosis;
KW cardiac hypertrophy; immune disorder; inflammation; autoimmune disease;
KW arthritis; graft rejection; inflammatory bowel disease.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH MISC-difference 2 /note="Encoded by AC"
FT
XX
XX US631424-B1.
XX
XX 18-DEC-2001.
XX
XX 15-JUN-2000; 2000US-0594669.
XX
XX 20-APR-1999; 99US-0295612.
XX
XX 18-MAY-1999; 99US-0314464.
XX
XX (CYTO-) CYTOKINETICS INC.
XX
XX Beraud C, Sakowicz R;
XX
XX WPI; 2002-089075/12.
XX
XX N-PSDB; AAD24085.
XX
XX New human MCAK (mitotic centromere-associated kinesin) protein useful
PT in identifying agents for use in the treatment of cellular
PT proliferation disorders -
XX
XX Claim 1; Fig 12; 44pp; English.
XX
XX The invention relates to human MCAK (mitotic centromere-associated
CC kinesin) protein, and its fragments that comprise a motor domain and
CC directly or indirectly produce ADP or phosphate. The MCAK enzyme and
CC its fragments are used in methods to identify compounds that modulate
CC their activity. Modulators of MCAK are useful as therapeutic agents
CC for treating cellular proliferation disorders such as cancer,
CC hyperplasias, restenosis, cardiac hypertrophy, immune disorders,
CC inflammation, autoimmune disease, arthritis, graft rejection,
CC inflammatory bowel disease. The present sequence is a human MCAK
CC fragment.
XX
XX Sequence 405 AA;
SQ
Query Match 11.0%; Score 752.5; DB 23; Length 405;
Best Local Similarity 42.1%; Pred. No. 1.1e-34;
Matches 1/4; Conservative 75; Mismatches 143; Indels 21; Gaps 10;
QY 4 KIKVVRKRLSELEKKKSDIITVKNCTLYIDPRYKVDWTKYIERHEFIVDKVPDD 63
DB 3 RLCVCVRKPLNKQELAKKEIDVISIPSKCLLVHPEPKLVLDITKYLENOAFEPDAFDE 62
QY 64 TVDNFTVYENTIKPLIIDYENGCVCSFAYGGTSGSKTYTMGS-QPYGQSDTPGIFQY 122
DB 63 TASNVEVYFTAPLAVQITIFEGG-KATCFAYGGTSGSKTHWGDSGKAQNNSKGIYAM 121
QY 123 AAGDITFTPLN--IYDQNTKGIISFYEYICGLYLLQKKRVALLENGKKEVYVVKDK 180
DB 122 ASRDVFLKNQPCYRKLGLE-VYVTFEPIYNGKLFDLNNKAKLRYLEBQKQOVVVGQ 180

```

```

QY 181 ILRLTKBELILKMD-GVLLRKIGVNSQNDSSRSHAILNIDKINNTSLGIAFID 239
DB 181 EHVNSADD-VIKIMDSACRISGOTFANSNSRSHAFQIILR--AGRNHGKSLVD 237
QY 240 LAGSERGADTVSONKOTQDGANINSLALKECTIPAMDSKXHIPRSELTKVLRDIF 239
DB 238 LAGNERGADTSSADROTRMEGAIEINKSLALKECTIPALGQNRKATPFRSKLTQVLRDSF 297
QY 300 VCK-SKSIIMIANISPTISCEQTLNTRYSRVKPKNKSTCINEEDDINTERISTLDSK 358
DB 298 IGENSRKCMATISPEISCEYTLNTRVADRVKELSPHSGSGEQ-----LQWE 348
QY 359 GSEMNASTIENVIKSNHLLSNNNNKIRGKINDKIEENNILKNKSPDKPRE 411
DB 349 TERMEACS--NGALLPGNLKSEBEELSSQMSFRNEMTQIRLEEKAMELKE 399
RESULT 14
AAE14503
ID AAE14503 standard; Protein; 434 AA.
XX
XX AAE14503;
AC
XX
XX 09-APR-2002 (first entry)
XX
XX Human mitotic centromere-associated kinesin protein fragment #5.
DE
XX
XX Human; MCAK; mitotic centromere-associated kinesin; motor domain;
KW cellular proliferation disorder; cancer; hyperplasia; restenosis;
KW cardiac hypertrophy; immune disorder; inflammation; autoimmune disease;
KW arthritis; graft rejection; inflammatory bowel disease.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH MISC-difference 2 /note="Encoded by CA"
FT
XX
XX US631424-B1.
XX
XX 18-DEC-2001.
XX
XX 15-JUN-2000; 2000US-0594669.
XX
XX 20-APR-1999; 99US-0295612.
XX
XX 18-MAY-1999; 99US-0314464.
XX
XX (CYTO-) CYTOKINETICS INC.
XX
XX Beraud C, Sakowicz R;
XX
XX WPI; 2002-089075/12.
XX
XX N-PSDB; AAD24084.
XX
XX New human MCAK (mitotic centromere-associated kinesin) protein useful
PT in identifying agents for use in the treatment of cellular
PT proliferation disorders -
XX
XX Claim 1; Fig 10; 44pp; English.
XX
XX The invention relates to human MCAK (mitotic centromere-associated
CC kinesin) protein, and its fragments that comprise a motor domain and
CC directly or indirectly produce ADP or phosphate. The MCAK enzyme and
CC its fragments are used in methods to identify compounds that modulate
CC their activity. Modulators of MCAK are useful as therapeutic agents
CC for treating cellular proliferation disorders such as cancer,
CC hyperplasias, restenosis, cardiac hypertrophy, immune disorders,
CC inflammation, autoimmune disease, arthritis, graft rejection,
CC inflammatory bowel disease. The present sequence is a human MCAK
CC fragment.
XX
XX Sequence 434 AA;
SQ

```


PR 20-APR-1999; 99US-0295612.
 PR 18-MAY-1999; 99US-0314464.
 XX
 PA (CYTO-) CYTOKINETICS INC.
 XX
 XX Berard C, Sakowicz R;
 PI
 DR MPI; 2002-089075/12.
 DR N-PSDB; AAD24086.
 XX
 XX
 PT New human MCAK (mitotic centromere-associated kinesin) protein useful
 PT in identifying agents for use in the treatment of cellular
 PT proliferation disorders -
 XX
 XX
 PS Claim 1; Fig 14; 44pp; English.
 XX
 CC The invention relates to human MCAK (mitotic centromere-associated
 CC kinesin) protein, and its fragments that comprise a motor domain and
 CC directly or indirectly produce ADP or phosphate. The MCAK enzyme and
 CC its fragments are used in methods to identify compounds that modulate
 CC their activity. Modulators of MCAK are useful as therapeutic agents
 CC for treating cellular proliferation disorders such as cancer,
 CC hyperplasia, restenosis, cardiac hypertrophy, immune disorders,
 CC inflammation, autoimmune disease, arthritis, graft rejection,
 CC inflammatory bowel disease. The present sequence is a human MCAK
 CC fragment.
 CC
 XX
 XX
 SO Sequence 723 AA;
 Query Match 11.0%; Score 752.5; DB 23; Length 723;
 Best Local Similarity 42.1%; Pred. No. 2.1e-34;
 Matches 174; Conservative 75; Mismatches 143; Indels 21; Gaps 10;
 QY 4 KIKVVRKRPPLSELEKKKDSIIIVKNNCTLYIDPRYKVMTKYIERHEFIVDAVPD 63
 DB 256 RLCVCRKRRLPKQELAKKEIDVISPSKCLLVHPEPKLVDTKYLENQAFCEFDPAFDE 315
 QY 64 TYDNFTVYENTIKPLIIDYENGCVSCPAVGGTSGKTYTLMGS-QPYGSDTPGIFQY 122
 DB 316 TASNEVVYFTARPLVQTIPEGG-KATCPAYGOTSGKTHMGDLSGKQANSKGIYAM 374
 QY 123 AAGDIFTFLN-IYDKNTKGIFISFEIYCGKLYDLQRRKVAALENGKEVVKDK 180
 DB 375 ASRDVFLKNQPCYRKLGLE-VVTFPEIYNGKLPDLNKKAKRLVLEDGKQOVVGLQ 433
 QY 181 IIRVLTKBELILKMD-GYLARKIGVNSQNDSSRSHALINDKINKNTSGKIAFID 239
 DB 434 EHLVNSADD-VIKIMDGSACTSGQTFANSSSRSHACFOILR-AKGRMHGKESLVD 490
 QY 240 LAGSERGADTVSONKOTOTDGANINNSIALKECIRAMDSKKNHIFPDSLETKVLRDIF 299
 DB 491 LAGNERGADTVSSADRTKMEGAELNKSLLALKECIRALQONKATHPFRESKLTQVLRDSF 550
 QY 300 VGR-SKSIIMIANISPTISCCCEQTLNLTLYSSRVKNFKNKSTCINEEDDTNTERISILDSK 358
 DB 551 IGENSRICMIATISPGISSCEYTLNLTLYADRVKELSPHSGPSGEG-----LIQME 601
 QY 359 GSEMASSIEVNVIKSNHLSNNNNKINRKGKINDKIERNNILKNSPKDKPRE 411
 DB 602 TEEMECAS--NGALIPGNLSKEEELSSQMSFNEAMTQIRELEKAMELKE 652
 Db
 RESULT 17
 ID AAE14506 standard; Protein; 725 AA.
 XX
 XX AAE14506;
 AC
 XX
 DT 09-APR-2002 (first entry)
 DE Human mitotic centromere-associated kinesin protein.
 XX
 XX Human; MCAK; mitotic centromere-associated kinesin; motor domain;

KW cellular proliferation disorder; cancer; hyperplasia; restenosis;
 KW cardiac hypertrophy; immune disorder; inflammation; autoimmune disease;
 KW arthritis; graft rejection; inflammatory bowel disease.
 XX
 OS Homo sapiens.
 XX
 XX US631424-B1.
 PN
 XX
 PD 18-DEC-2001.
 XX
 PF 15-JUN-2000; 2000US-0594669.
 XX
 XX
 XX 20-APR-1999; 99US-0295612.
 PR 18-MAY-1999; 99US-0314464.
 XX
 XX
 PA (CYTO-) CYTOKINETICS INC.
 XX
 XX Berard C, Sakowicz R;
 PI
 DR MPI; 2002-089075/12.
 DR N-PSDB; AAD24087.
 XX
 XX
 PT New human MCAK (mitotic centromere-associated kinesin) protein useful
 PT in identifying agents for use in the treatment of cellular
 PT proliferation disorders -
 XX
 XX
 PS Disclosure; Fig 16; 44pp; English.
 XX
 CC The invention relates to human MCAK (mitotic centromere-associated
 CC kinesin) protein, and its fragments that comprise a motor domain and
 CC directly or indirectly produce ADP or phosphate. The MCAK enzyme and
 CC its fragments are used in methods to identify compounds that modulate
 CC their activity. Modulators of MCAK are useful as therapeutic agents
 CC for treating cellular proliferation disorders such as cancer,
 CC hyperplasia, restenosis, cardiac hypertrophy, immune disorders,
 CC inflammation, autoimmune disease, arthritis, graft rejection,
 CC inflammatory bowel disease. The present sequence is human MCAK protein.
 CC
 XX
 XX
 SO Sequence 725 AA;
 Query Match 11.0%; Score 750.5; DB 23; Length 725;
 Best Local Similarity 41.9%; Pred. No. 2.7e-34;
 Matches 173; Conservative 76; Mismatches 143; Indels 21; Gaps 10;
 QY 4 KIKVVRKRPPLSELEKKKDSIIIVKNNCTLYIDPRYKVMTKYIERHEFIVDAVPD 63
 DB 258 RLCVCRKRRLPKQELAKKEIDVISPSKCLLVHPEPKLVDTKYLENQAFCEFDPAFDE 317
 QY 64 TYDNFTVYENTIKPLIIDYENGCVSCPAVGGTSGKTYTLMGS-QPYGSDTPGIFQY 122
 DB 318 TASNEVVYFTARPLVQTIPEGG-KATCPAYGOTSGKTHMGDLSGKQANSKGIYAM 376
 QY 123 AAGDIFTFLN-IYDKNTKGIFISFEIYCGKLYDLQRRKVAALENGKEVVKDK 180
 DB 377 ASRDVFLKNQPCYRKLGLE-VVTFPEIYNGKLPDLNKKAKRLVLEDGKQOVVGLQ 435
 QY 181 IIRVLTKBELILKMD-GYLARKIGVNSQNDSSRSHALINDKINKNTSGKIAFID 239
 DB 436 EHLVNSADD-VIKIMDGSACTSGQTFANSSSRSHACFOILR-AKGRMHGKESLVD 492
 QY 240 LAGSERGADTVSONKOTOTDGANINNSIALKECIRAMDSKKNHIFPDSLETKVLRDIF 299
 DB 493 LAGNERGADTVSSADRTKMEGAELNKSLLALKECIRALQONKATHPFRESKLTQVLRDSF 552
 QY 300 VGR-SKSIIMIANISPTISCCCEQTLNLTLYSSRVKNFKNKSTCINEEDDTNTERISILDSK 358
 DB 553 IGENSRICMIATISPGISSCEYTLNLTLYADRVKELSPHSGPSGEG-----LIQME 603
 QY 359 GSEMASSIEVNVIKSNHLSNNNNKINRKGKINDKIERNNILKNSPKDKPRE 411
 DB 604 TEEMECAS--NGALIPGNLSKEEELSSQMSFNEAMTQIRELEKAMELKE 654

RESULT 18
AAB18301
ID AAB18301 standard; Protein, 1817 AA.
XX AAB18301;
AC
XX
DT 07-NOV-2000 (first entry)
XX
DE Plasmodium falciparum chromosome 2 related protein SEQ ID NO:159.
XX
XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
KM antimalarial; malaria; protozoacide; infection; insecticide.
XX
OS Plasmodium falciparum.
XX
PN W0200025728-A2.
XX
PD 11-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US26796.
XX
PR 05-NOV-1998; 98US-0107131.
XX
PA (HOFF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
XX
PI Hoffman S, Carucci D, Gardner M, Venter JC,
XX
XX WPI: 2000-365347/31.
PT Proteins encoded by chromosome 2 of the human malarial parasite.
PT Plasmodium falciparum, useful as antimalarial vaccines and in the
PT diagnosis of P.falciparum infection -
XX
XX Disclosure: Page 374-379; 577pp; English.
XX
XX The present invention describes proteins and their fragments (I) encoded
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC Also described are: (I) nucleotide sequences (II) encoding (I); and (2)
CC vaccines against P. falciparum infection comprising (I) or (II).
CC (I) and (II) are useful for the development of vaccines against
CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal
CC antibody raised to immunogens comprising the sequences of (I), are
CC useful in the detection of infection with P. falciparum. Furthermore,
CC (I) (especially when they are ritins or secreted or membrane proteins)
CC can aid the identification of drugs to treat or prevent P. falciparum
CC infection, or they can be used to identify drug resistance in
CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
CC subsequent identification of proteins encoded by it will help to expand
CC our understanding of parasite biology, a process hampered by the
CC complexity of the parasite lifecycle, and provide new targets for
CC vaccine and drug development. Parasite resistance to drugs and mosquito
CC resistance to insecticides have led to a resurgence of malaria in many
CC parts of the world, and there is a pressing need for vaccines and new
CC drugs. AAT70078 to AAT70287 and AAB18144 to AAB18352 represent nucleotide
CC and protein sequences given in the present invention, but which are not
CC specifically mentioned within the specification.
XX
SQ Sequence 1817 AA;
Query Match 10.9%; Score 748.5; DB 21; Length 1817;
Best Local Similarity 23.6%; Pred. No. 9,8e-34;
Matches 356; Conservative 197; Mismatches 410; Indels 543; Gaps 69;
QY 79 IIDLYENCVCSCFAYGOTGSGKTYTLMGSGPYGSGDTPGIF-----OYAAGDIFTFANI 133
D 391 IPELFTBECYLFLEPNEKIKSNIIY-----IPNIFLKKLKYIISNDV---I 435
QY 134 YDKONTKGIFFISFYIYCGK-----LYDLQKRMVAALNGKKEVVVKDLK--- 180
D 436 HINDRTLRNRYIKFISYHCSNNIYVAILPNDLYITLHKXIFI-----NNYDMLIKHYKSTA 490

QY 181 --ILRVLTKEELLKMGIDVLRK-IGVNSQNDSSSRSHAIIIDLK-----DI 226
D 491 DHILYLTNKE-NLNTYNTILLNNYSLYVKNKENNERQSLNMLKXIPSHLYKQHTT 549
QY 227 NKQTSIACKIAFIDLAGSGADTVSQNKQOTQTCANINRSLALKECIRAMDSDKHIIF 286
D 550 NKQV-----DPQGHKDKD-----SYDNTVEQMKNNKNTKYP 584
QY 287 RDSLETLVADIFVGSKSIIMANISPTISCCQTLNTLAYSRSVKNFKKSTCTINEDD 346
D 585 NEYTTTHLDN-----NYEONLSFQK-----DD 609
QY 347 TTERISILDSKSGEMNASSIENVVIKSNHLNNNNNKINRGKINDKIRNNILK---- 402
D 610 TNINNIIDL-HKRGQIYEYKEN---ESSDIFPDSYRKIKKEK-----KKNIYKEDH 660
QY 403 --NKSFPKPREGFTSTFGKYSLSNDIDKIKKKKKGLINYKSTLYN---DNTINKKINN 457
D 661 PLNKEKKKKKKFFYINYEKDDKDNDL-----YNNIYSKULENIQKKNYNN 709
QY 458 NNNNNNDNNNDNNNNNNNNNDSSSMVNMIMNINNNINNNNNNNNNNNNNNNNN 517
D 710 NNN 756
QY 518 HLPQNYAFTDTSDFSSL-----DDMCHLNNDKSIPL-LHKQLRDNK-----LKN 564
D 757 NIPCTNINLSLYKEXELPLFYAYDKIQYTYEELSKYKISITKIDRNIKMLFKFLKN 816
QY 565 RSS-----CDNINKKKNNLHARHVSGLTMEFSPQOKONT-----FFSS----- 608
D 817 YNNNENTYVDNIIS-KENIFHLLA-SMKNKVT-----NKTTHKDIYQFIHSHWYHK 866
QY 609 --NINK-----MEDTPDIIYESRNVSNMGNVLLGKNTKTHDISTRDEN-----HND 656
D 867 LADQNKHSFODDKYLLINLVEKHKVQ-----NMTHTIHKEDKKGDHLMENN 917
QY 657 NKINNVINIINNSVNS-----INNS-----NNNSINNS-NMNS 690
D 918 NMLNNMSLNNSNLSNLSIPLNNSIPLNNSIPLNNSIPLNNSIPLNNSIPLNNS 977
QY 691 -----NSIYSNYSNOSISDVQIR-YVEMQDS----- 718
D 978 CISLYNSISYSNKSFTNHLNNTYDTCFIONNYSNOOVYNNKKNMEHYNEKUL 1037
QY 719 -----NKNNDNIF----- 727
D 1038 FIYIYLEDKNYFLNVNNNIFPKKNNTNTEFTYQINILSKLYFINVYTLISSNYK 1097
QY 728 -DAISCDNNYPTNTNN 786
D 1098 ABEIKTDNKC-NINNN 1149
QY 787 DNNKNTS-----NIONINTNK-----NNQGVNYSNFCF----- 817
D 1150 INHFTTTFYEVATYLLKNIYAIHISKEFYIFVALSKEFLNYSXOOSTNRKENSIMENV 1209
QY 818 -----YLANDKNTLIDLNNKEQDKNIGHGDDNIIQI-----RND 853
D 1210 LYLIYIIRKQYEHVKSILYDKSNENYFRNENKDIKMEINTMLYNIILNNSFEDHDF 1269
QY 854 EKKKKTFFYNNNNIIVIVNNNMGNNSPRMKYGLCGSHTSIDNMKNEMKNEMKNEMKD 913
D 1270 MTLQKNEDNNKMIIDININVDNIN-----DLISHCDNNKED----- 1309
QY 914 NHIKSNNSSSSSSSNNNINNNINDDTPQNDYCHANDNFTTIRKKNNTINSIYOND 973
D 1310 -----TSLHNKLYNGLHFLMFLANNYLDTKHKFI-----NHFSSFYTNK- 1352
QY 974 IITNTINSUNDYMSLTLL-----FKKYTYPTLSTNEDIYV-----KEMEGKHRLDDQD 1023
D 1353 -IIPNMGMHYTELTYLHKHKIKKPKFY-----IYNGDLKKKSYLVHVK----- 1398


```

QY 1024 KYDDNNNNVNDNNKNNVNDNNVNDNNVNDNN--DKNNVNDNNVNDNDDVDVFNHKN 1081
DB 1399 KLYNSYIKSYNNKKK-----NNVNGDVNNFMFKYNI-----YDNIDYFIKK 1443
QY 1082 -----FNNNEVLSYFQKNDVTIINNCLNSLSDISMYDDTFEILNLLSKYKAE----KD 1132
DB 1444 KMLFCYTNHSLSLYF-----TYIYS--LNKPYCYTLTYN-----ISKCPYKININENIHFKN 1493
QY 1133 NVIKKYNINEDIKNMSLEIDKTAQSIY---EKRRV-----LTKLLLFKKKVDVQINN 1183
DB 1494 KILFFIFQCKKIYIKFFFLVQSIFFSSSEFQKVGKLSYILSNILLVKSRRKKN 1553
QY 1184 ETSDLKDLVMGCI CANNPDQFHYAVYSKLEKDIINLMKQIWE---SENRLLYQ 1239
DB 1554 KKKKIYKN--ISKIYSNN-----EFINNKKIKKHITVNNSSMKLFIQNK 1597
QY 1240 FLVVEYQ-----NKSANSVLLNVSNNNGDIILANKLVQDNKN 1278
DB 1598 LNLTIQNNVYFPMDLFISSNLSHETELINLKEQNIILNNNNNNN-----NNNNN 1647
QY 1279 SMDHNN 1284
DB 1648 NKKHNN 1653

RESULT 19
AAE14501
ID AAE14501 standard; Protein; 362 AA.
AC AAE14501;
XX
XX
DT 09-APR-2002 (first entry)
XX
DE Human mitotic centromere-associated kinesin protein fragment #3.
XX
KW Human; MCAK; mitotic centromere-associated kinesin; motor domain;
KW cellular proliferation disorder; cancer; hyperplasia; restenosis;
KW cardiac hypertrophy; immune disorder; inflammation; autoimmune disease;
KW arthritis; graft rejection; inflammatory bowel disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 2 /note= "Encoded by AC"
XX
XX
PN US6331424-B1.
XX
PD 18-DEC-2001.
XX
PF 15-JUN-2000; 2000US-0594669.
XX
PR 20-APR-1999; 99US-0295612.
PR 18-MAY-1999; 99US-0314464.
XX
PA (CYTO-) CYTOKINETICS INC.
XX
PI Beraud C, Sakowicz R;
XX
DR WPI; 2002-089075/12.
DR N-PSDB; AAD24082.
XX
PT New human MCAK (mitotic centromere-associated kinesin) protein useful
PT in identifying agents for use in the treatment of cellular
PT proliferation disorders -
XX
XX
XX Claim 1; Fig 6; 44p; English.
XX
XX The invention relates to human MCAK (mitotic centromere-associated
XX kinesin) protein, and its fragments that comprise a motor domain and
XX directly or indirectly produce ADP or phosphate. The MCAK enzyme and
XX its fragments are used in methods to identify compounds that modulate
XX their activity. Modulators of MCAK are useful as therapeutic agents
XX

```

```

CC for treating cellular proliferation disorders such as cancer,
CC hyperplasia, restenosis, cardiac hypertrophy, immune disorders,
CC inflammation, autoimmune disease, arthritis, graft rejection,
CC inflammatory bowel disease. The present sequence is a human MCAK
CC fragment.
XX
SQ Sequence 362 AA;
XX
Query Match 10.9%; Score 745.5; DB 23; Length 362;
Best Local Similarity 45.7%; Pred. No. 2,3e-34;
Matches 188; Conservative 65; Mismatches 116; Indels 19; Gaps 9;
QY 4 KIKVVRKRPKLELEKKKSDIIVKNNCTIYIDPRYKVMYTKYIERHFIYDKVFPD 63
DB 3 RICVCVRKRPKNQBLAKKEIDVISIPSKCLLVHBPVKLVDTKYLENOAFCEFPADF 62
QY 64 TYDNFTYENTIKPLIIDYENGCVSCFAYGQTSQKTYTMLGS-QPYGQSTPGLFQY 122
DB 63 TASNVEVYFRTARPLVQITFEFG-KATCPAYQQTSGKHTMGDLSGKAQNASKGIYAM 121
QY 123 AAGDIFTPLN--IYDKNTKGIFISFYECGKLYDLQKRWVALENGKKEVVVKDK 180
DB 122 ASHDFVLKNQPCYRKGLGE-VYTFEITNGKLPDLNKKAKRLVLEQKQVQVVGID 180
QY 181 ILRLTKRELILKMD-GVILRKIGVNSQNDSSRSRSHALINIDLKINKNTSLGTAFTD 239
DB 181 EHLVNSADD-VIKMIDMSACRTSGQTFANSSSRSHACFOQILR--AAGRMHGRKSLVD 237
QY 240 LAGSEFGADTVSONKQOTDGANINRSLALKECTIRAMSDKNHIFPDSLTQVLRDIF 299
DB 238 LAGNERGADTSSADROTRMEGABINKSLALKECTIRALQONKAHTPFRESKLTQVLRDSF 297
QY 300 VCK-SKSIWIANISPTSCCEOTLNTLRYSSRVKPKNKNSTCINEEDDTNTERISLDSK 358
DB 298 IGENSKTMTATISPEISCEYTLNTRYADRVKELSPHSGPSGQ-----LIQME 348
QY 359 GSEMNASS 366
DB 349 TEMEMACS 356

RESULT 20
AAE14500
ID AAE14500 standard; Protein; 391 AA.
AC AAE14500;
XX
XX
DT 09-APR-2002 (first entry)
XX
DE Human mitotic centromere-associated kinesin protein fragment #2.
XX
KW Human; MCAK; mitotic centromere-associated kinesin; motor domain;
KW cellular proliferation disorder; cancer; hyperplasia; restenosis;
KW cardiac hypertrophy; immune disorder; inflammation; autoimmune disease;
KW arthritis; graft rejection; inflammatory bowel disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 2 /note= "Encoded by CA"
XX
XX
PN US6331424-B1.
XX
PD 18-DEC-2001.
XX
PF 15-JUN-2000; 2000US-0594669.
XX
PR 20-APR-1999; 99US-0295612.
PR 18-MAY-1999; 99US-0314464.
XX
PA (CYTO-) CYTOKINETICS INC.
XX

```


Query Match	10.9%	Score 745.5	DB 23	Length 391
Best Local Similarity	45.7%	Pred. No. 2.6e-34		
Matches 168	Conservative 65	Mismatches 116	Indels 19	Gaps 9
QY	4 KIKVVRKRKPELEKKKDSIIIVKNNCTLYIDPRYKVMYKIERHFIIVDKVPED	63		
DB	32 RLCVCRKRPPLKQELAKKEIDVISIPSKCLLVNHPKLVLTLYLENQAFCPFAFDE	91		
QY	64 IVDNFTYVNTIKPLIIDLYENGCVCSCPAYQQTSGKTYTMGS-QPYGSDTPGIFQY	122		
DB	92 TASNENVVYFTARPLVQTIFEGG-KATCPAYQQTSGKTHWGGDLSGKQVASKGIYAM	150		
QY	123 AAGDIFTFLN--IYDNDTKGFIISFYELYCKLVYLLQKKRMVALENGKEVYVYKDX	180		
DB	151 ASRDVFLKNOCPYKRLGLE-VYVTFEFLYNGFLFDLKKKALRYLLEGGKQOVVGVGQ	209		
QY	181 ILRVLTKEELIKMID-GVLLRKIRIGVNSQNDSSSRSHALINTDLKINKNTSGIKAFID	239		
DB	210 EHLVNSADD-VIKMIDMGACRPSGQTFPANSNSRSHACFOILLR--AGRGHGRFSLV	266		
QY	240 IAGSERGADTVSONKQTOTGDGANINISLLALKECTIRAMDSKNHPIFRDSELTXYLRDIF	299		
DB	267 LAGNERGADTVSADRGTRMEGAENKSLALKECIPALGQNKATPTFRESKLTQVLRDSF	326		
QY	300 VGR-SKSTIMIANISPTISCEQTLNLTFRSSRYKRNKNSKSTCINEDDPTIRISILDSK	358		
DB	327 IGENNSTKMIATISPEISSCGYTLNLTFRYADRYKELSPHSGPGEQ-----LIQWE	377		
QY	359 GSEMNASS 366			
DB	378 TEEMERAS 385			
RESULT 21				
AAEL14499				
ID	AAEL14499 standard; Protein; 430 AA.			
AC	AAEL14499;			
XX				
DT	09-APR-2002 (first entry)			
XX				
DE	Human mitotic centromere-associated kinesin protein fragment #1.			
XX				
KW	Human; MCAK; mitotic centromere-associated kinesin; motor domain;			
KW	cellular proliferation disorder; cancer; hyperplasia; resenosis;			
KW	cardiac hypertrophy; immune disorder; inflammation; autoimmune disease;			
XX	arthritis; graft rejection; inflammatory bowel disease.			
OS	Homo sapiens.			

[illegible]

[illegible]

Dd		46	VDLAIGNERGADQGSADRQRIRIEGAELINSLALKECICRALRQSGSHLPFRSKUTQVLRD	528
Oy		298	IFPG--KKSIMTAMISPTSCCEQTALTLYSSAVKNF-----KKSTCI	341
Dd		529	SFGYKKKKTKTCMTAMISPSMSCVENLTNTLRYADRVKELAKEDHQLQSVEGDGCKSPDL	588
Oy		342	NEE-----DDTTERISILDSKGSEMANSSIENVIVISSNHLSNNNNKI	386
Dd		589	NESSEPEEMADEBGEDEEPEDENOHLTTISSEASVYNMSTD--MSGNHTL--NIIGPS	643
Oy		387	NRSKINDKIERNNILNKNSFDRRGFFSTEGKYSSLNDIDIKIKKKGGIINYKSTL--	444
Dd		644	RNVDIGEAEGHALVEN-----LETFANHRQLKTDEKIEIQYONGBSALMKULAVNR	698
Oy		445	-----YNDNTINKKHNN	457
Dd		699	TRDVTHNYVTQKLKKEENQN	718
RESULT 23				
AAM39258	ID	AAM39258	standard; Protein; 728 AA.	
XX	AC	AAM39258;		
XX	DT	22-OCT-2001	(first entry)	
XX	DE	Human polypeptide SEQ ID NO 2403.		
XX	KW	Human; nocrotropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KW	KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;		
XX	OS	leukaemia.		
XX	XX	Homo sapiens.		
XX	PX	WO200153112-A1.		
PD	XX	26-JUN-2001.		
XX	PF	26-DEC-2000; 200OWO-US34263.		
XX	PR	21-JAN-2000; 200OUS-0488725.		
PR	PR	25-APR-2000; 200OUS-0552317.		
PR	PR	09-JUL-2000; 200OUS-0598042.		
PR	PR	19-JUL-2000; 200OUS-0620312.		
PR	PR	03-AUG-2000; 200OUS-0653450.		
PR	PR	14-SEP-2000; 200OUS-0662191.		
PR	PR	19-OCT-2000; 200OUS-0693036.		
XX	PR	29-NOV-2000; 200OUS-0727344.		
PA	XX	(HYSEQ-) HYSEQ INC.		
XX	PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,		
PI	PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J,		
PI	PI	Zhao QA, Zhou P, Goodrich R, Dimanac RT;		
XX	XX	WPI; 2001-442253/47.		
DR	XX	N-PSDB; AAI58414.		
XX	PT	Novel nucleic acids and polypeptides, useful for treating disorders		
XX	XX	such as central nervous system injuries -		
PS	XX	Example 4; SEQ ID NO 2403; 10078bp; English.		
CC	XX	The invention relates to human nucleic acids (AA157798-AA161369) and		
CC	CC	the encoded polypeptides (AAM38642-AAM42213) with nocrotropic,		
CC	CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
CC	CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	CC	of the invention may be used to treat diseases of the peripheral nervous		
CC	CC	system, such as peripheral nervous injuries, peripheral neuropathy and		

XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA38642-AA42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

SQ Sequence 690 AA;

Query Match 10.7%; Score 732.5; DB 22; Length 690;

Best Local Similarity 45.5%; Pred. No. 2.6e-33; Mismatches 111; Indels 25; Gaps 9;

Matches 166; Conservative 63; Mismatches 111; Indels 25; Gaps 9;

QY 4 KIKVVRKRPRLSELEKKKKSDITVKNKCTLYIDEPKRVDMTKYIERHFEIVDKVDD 63
 DB 207 RICVCRKRPPLNKETOMKDLVITIPSKDVVMHPEKQKVDLRYLENQTFRFDVAFDD 266
 QY 64 TVDNFTYENTIKPLIIDLYENGVCSCFAYGQSGKTYTMLS-QPYGSDTPGIFQY 122
 DB 267 SAPNEMVYRFTARPLVETIFERG-MATCFAYGQSGKTHMGDFSGKNDSCSGIYAL 325
 QY 123 AAGDIFFTFLNYDKONTK---GIFISFYEIYCGKLYDLQKRMVAALENGKEVYVKD 178
 DB 326 AARDVFLMLK--KPNYKLELYATFPEIYSGKVFLLNRKTKLRYLEBGKQOVVVG 382
 QY 179 LKILRVLTKEELIKMID-GVLLRKIGVNSQNDSSSHALINLDIDKNNTSLGKIAF 237
 DB 383 LOBEREVKVED-VLKLDIGNSCSTSGQTSANNAHSSSHAVFQILR-RKGKLGKRS 439
 QY 238 IDLGSERGADTVSONKOTQTDGANINSLALKECIRAMDSKRNHIFPRDSELTXYLD 297
 DB 440 IDLAGNERGADTSSADQTRLEGAINKSLALKECIRALGRNKPFPFASKULTQVLRD 499
 QY 298 IFVKG-SKSIIMIANISPTISCEQTLNLTARYSSRVKNFKNSGTCT-----NEED 345
 DB 500 SPFGNSRFTCMATISPGMASCENTLNTLRYANRVKELTVDEPTAAGDVPRIMHPPNQID 559
 QY 346 DTNTE 350
 DB 560 DLETO 564

RESULT 26

AA41043 ID AA41043 standard; Protein; 694 AA.

AC AA41043;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 5974.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

OS Homo sapiens.

XX WO20015312-A1.

PN

XX 26-JUL-2001.
 PD 26-DEC-2000; 2000MO-US34263.
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

DR N-PSDB: AA160199.

PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -

PS Example 2; SEQ ID NO 5974; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA38642-AA42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

SQ Sequence 694 AA;

Query Match 10.7%; Score 732.5; DB 22; Length 694;

Best Local Similarity 45.5%; Pred. No. 2.6e-33; Mismatches 111; Indels 25; Gaps 9;

Matches 166; Conservative 63; Mismatches 111; Indels 25; Gaps 9;

QY 4 KIKVVRKRPRLSELEKKKKSDITVKNKCTLYIDEPKRVDMTKYIERHFEIVDKVDD 63
 DB 211 RICVCRKRPPLNKETOMKDLVITIPSKDVVMHPEKQKVDLRYLENQTFRFDVAFDD 270
 QY 64 TVDNFTYENTIKPLIIDLYENGVCSCFAYGQSGKTYTMLS-QPYGSDTPGIFQY 122
 DB 271 SAPNEMVYRFTARPLVETIFERG-MATCFAYGQSGKTHMGDFSGKNDSCSGIYAL 329
 QY 123 AAGDIFFTFLNYDKONTK---GIFISFYEIYCGKLYDLQKRMVAALENGKEVYVKD 178
 DB 330 AARDVFLMLK--KPNYKLELYATFPEIYSGKVFLLNRKTKLRYLEBGKQOVVVG 386
 QY 179 LKILRVLTKEELIKMID-GVLLRKIGVNSQNDSSSHALINLDIDKNNTSLGKIAF 237
 DB 387 LOBEREVKVED-VLKLDIGNSCSTSGQTSANNAHSSSHAVFQILR-RKGKLGKRS 443
 QY 238 IDLGSERGADTVSONKOTQTDGANINSLALKECIRAMDSKRNHIFPRDSELTXYLD 297
 DB 444 IDLAGNERGADTSSADQTRLEGAINKSLALKECIRALGRNKPFPFASKULTQVLRD 503
 QY 298 IFVKG-SKSIIMIANISPTISCEQTLNLTARYSSRVKNFKNSGTCT-----NEED 345
 DB 560 DLETO 564

Db 504 SFIGENSRTCMATISPGMASCENTLNTLRVANKVELTVDPTAAGDVRPIMHHPNOID 563

QY 346 DTNTE 350

Db 564 DLETFQ 568

RESULT 27

AA41044 ID AAM41044 standard; Protein; 694 AA.

XX AAM41044;

AC AAM41044;

XX 22-OCT-2001 (first entry)

DT Human polypeptide SEQ ID NO 5975.

DE Human, noctropis; immunosuppressant; cytostatic; gene therapy; cancer;

XX peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KM leukaemia.

OS Homo sapiens.

XX MO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000MO-US34263.

XX 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

DR N-PSDB; AAI60200.

XX Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

XX Example 2; SEQ ID NO 5975; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and

CC the encoded polypeptides (AAM38642-AAM42213) with noctropic.

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemia and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

XX Sequence 694 AA;

XX SQ

Query Match 10.7%; Score 732.5; DB 22; Length 694;

Best Local Similarity 45.5%; Pred. No. 2,6e-33;

.Matches 166; Conservative 63; Mismatches 111; Indels 25; Gaps 9;

QY 4 KIRVTVRKRPPLSELEKKKQSDIIVKNNCTLYIDPRYKVDMTKYIEHNEFIVDKVFPD 63

Db 211 RICVCVRKRPPLNKETQMDLDVITIPSKDVVWVHPKOKVDLTRYLENQTFRFDYAFDD 270

QY 64 TVDNFTVYENTTKPLIIDYENGCVCSFAYGQTSKTYTLMGS-QPYGQSDPTPGIFQY 122

Db 271 SAPNEVYRFTAPLVETIFERG-MATCFAYGQTSKTYTLMGSDPSGKXQDCSKITAYL 329

QY 123 AAGDIFTPLNIYDKDNTK----GIFSFYEIYCGKLYDLLQKKWVALENGKEVAVVD 178

Db 330 AARDVFLMK---KPYVKLELVQVATFEIYSGKVFDDLNNKTKLRVLEDDGQYQVVG 386

QY 179 LKILRVLTREELIKMID-GVLLRKIGVNSQNDSSRHAIINIDKINKTSLCKIAF 237

Db 387 LQREYKCYED-VLKLIDIGNSCRTSGQTSANAHSSRSYAVFOIILR--RKGLHGKFSL 443

QY 238 IDLAGSERGADTVSQKQKOTGDGANINRSLAKECIRAMDSDKXHIPRDSGLTVLVD 297

Db 444 IDLAGSERGADTVSSADROTLEGAEINSLALKECIRALGKXKHPPTPRASKLTVLVD 503

QY 298 IFVGK-SKSIMIANISPTISCEQTLNTRYSRVNFFXKSTCI-----NEED 345

Db 504 SFIGENSRTCMATISPGMASCENTLNTLRVANKVELTVDPTAAGDVRPIMHHPNOID 563

QY 346 DTNTE 350

Db 564 DLETFQ 568

RESULT 28

AA18199 ID AAB18199 standard; Protein; 1436 AA.

XX AAB18199;

XX 07-NOV-2000 (first entry)

DT Plasmodium falciparum chromosome 2 related protein SEQ ID NO:56.

XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;

KW antimalarial; malaria; protozoacide; infection; insecticide.

XX Plasmodium falciparum.

XX WO200025728-A2.

XX 11-MAY-2000.

XX 05-NOV-1999; 99WO-US26796.

XX 05-NOV-1998; 98US-0107131.

XX (HOFF/) HOFFMAN S.

PA (CARU/) CARUCCI D.

PA (GARD/) GARDNER M.

PA (VENT/) VENTER J C.

PI Hoffman S, Carucci D, Gardner M, Venter JC;

XX WPI; 2000-365347/31.

DR Protein encoded by chromosome 2 of the human malarial parasite,

PT Plasmodium falciparum, useful as antimalarial vaccines and in the

PT diagnosis of P.falciparum infection -

XX Disclosure, Page 133-137; 577bp; English.

XX The present invention describes proteins and their fragments (I) encoded


```

QY 689 NSNSIYKSNYNSQS-1SDVQIRYVN-EMDTSKNDNDIFPDAISCDNNMYPIITNNNN 746
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 380 SPENHILNMLNNSNSLPLSPPHINPHIVSNLNNNN-----NNINPNNNPNNSN 431
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 747 NNNNN---NNNNNIDVENYNNEDGTNNSKLYAYNSHNLFPDNNKNTSINQINTNKN 803
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 432 NSNNNVPNNNNINISPRGSNISPRSN-----GGSSTISPRNISNNNNIIN----- 478
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 804 NQDGNVYNSMNFCHYMLNDKNYLIDLNKQKDKNHGCD-----NNIIONNDFE 854
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 479 -----NNNNNLTTPPNSPRLE-NVNPNTSPRLATSLTLPVSSLT 522
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 855 KKKKTFFYNNNNIVIVNNNNNNNSPRMKYGLGSHTSIDNMKNEMKNEMKNEDKN 914
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 523 SSNNNNQSNNTTPSINNNGR-----GHCIQTISEILNKPVYVYNNNGNN----- 570
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 915 HIKSNNNSSSSSSNNNININDDTFQNDYCHNDNFTTIRKKNNTINSNIYQ---- 970
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 571 -----NNNNNTNSTTNNNITNNNNNN-----NNNNINNNVLTSPRK 607
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 971 ----NDIITYTNSLDYMSNTLHFKKXTTYPTLSTNEDIYKMEGKHRLDDQKYP 1026
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 608 RTKGNHS---KTNSLDPEFS-----SMNGGDSISGAGSGSLRRRRKDDND 652
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1027 DND-----NNNVNKNKNVNDNNVNDNNVNDNNVNDNNVNDNNVNDND 1073
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 653 ENDGNSNNTNNSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 712
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1074 VDFHNKFNENNNEYLSYFQKNVDTIINNCLNSLDISMYDTKEIILNNILSKYKAEKN 1133
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 713 NNNNNNNNNNNNN-----NNYNHGAITYMMSH-----NHQSIGSSSPKKN 752
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1134 VIKYINEDIKJMSLEIDKTAOSIYEKRVLTLLFKKRVNDVQINNETDLKOLY 1193
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 753 NFEPF---SRNCSLGMGRRAVAL-----ILGLF-----YVSSISILATLV 791
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1194 MCHICNNPDQHFYAVSRLEKDIINLIMLRQIWCESENRLLYQFLVVEYQNK----S 1249
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 792 LRSEENSIAADF-----ARVARD--RFTYLR-----IEPNRLIYITO 827
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1250 ANSVLLNV-SSNNGDITLANKKLVDON 1275
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 828 TSLSLTSEFPSTSEDQFVFPFSKLMSDN 854
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 32

AAB56496 standard; Protein; 410 AA.

AAB56496.

13-MAR-2001 (first entry)

Human prostate cancer antigen protein sequence SEQ ID NO:1074.

Human; prostate cancer; prostate cancer antigen; detection; diagnosis; neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular; vulnary; gastrointestinal; nephrotoxic; antineoplastic; gynaecological; antibacterial; gene therapy; neural; immune; reproductive; renal; gastrointestinal; pulmonary; cardiovascular; proliferative disorder; wound; infectious disease.

Homo sapiens.

WO200055174-A1.

21-SEP-2000.

08-MAR-2000; 2000MO-US05988.

12-MAR-1999; 99US-0124270.

XX

```

PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX Rosen CA, Ruben SM;
PI WPI; 2000-587513/55.
DR N-PSDB; AAF15699.
XX
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX
PS Claim 11; Page 1498-1499; 2338bp; English.
XX
CC AAF1566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytosolic,
CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
CC nephrotoxic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 410 AA;
XX
Query Match 9.8%; Score 672; DB 21; Length 410;
Best Local Similarity 46.6%; Pred. No. 3 7e-30;
Matches 149; Conservative 59; Mismatches 102; Indels 10; Gaps 8;
XX
QY 4 KIRVVVRKPLSELEKKKSDIIVKNNCTLYIDEPKRVMTXYIERHEPIVDKVPDD 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 57 RICVCVRKPLNKQELAKKEIDVISPSKCLLVHEPKLVDTXYLENQAFCEPAFDE 116
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 TYDNFTYENTIKPLIIDYENGCVSCFAYQGTGSGKTYTLMGS-QPGQSTPFGIFQY 122
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 117 TASNEVYRFTARPLVQITFEGG-KATCFAYGQGTGSKTHYMGDLSGAQNSKGIYAM 175
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 123 AAGDIFTPLN--LYDCDNTKGIFISFEYIYCGVLLKRRKVAALENGKKEVVVVKDK 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 176 AKRDVFLKNQPCYRKLGLF-VYVTFEYINGKLFPLNKKAKLRVLEGGQGVQVGLQ 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 ILRLTKKEIILKMD-GVILRKIGVNSQNDSSRSASHIINIDLDKINNTSISGTAIFD 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 235 EHLVNSAD-VIKMKMGSACTSGOTFANSSRSRHACFQIILR--AGRMHMKFSVLD 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 240 LAGSERGADTVSONKOTQTDGANINSLALKECIKRAMSDKNHIFPDSLTGVLDIF 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 292 LAGNERGAYTSSADROTRMEGAIEINKSLALKECIKALQNKVHTFPRESKLTQVLRDSF 351
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 300 VCK-SKSIIMIANISPTISCC 318
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 352 IGENSRCTMIATISPEISSC 371
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 33

AAB05593 standard; Protein; 604 AA.

AAB05593.

22-APR-2002 (first entry)

Mitotic centromere-associated kinesin NYD-KIF2 protein.

Human; mitotic centromere-associated kinesin; NYD-KIF2; human genome technology.

XX

OS Homo sapiens.
 PN CN1318561-A.
 XD 24-OCT-2001.
 XX 11-APR-2001; 2001CN-0113509.
 PF 11-APR-2001; 2001CN-0113509.
 PR 11-APR-2001; 2001CN-0113509.
 XX (UTNA-) UNIV NANTING MEDICAL.
 XX Sha J, Zhou Z, Li J;
 XX P1
 XX WPI: 2002-122927/17.
 DR N-PSDB; ABA934607.
 XX
 PT Human testicular mitosis dynein gene encoded protein for immunizing an
 XX animal and preparing monoclonal and polyclonal antibodies
 PS Disclosure; Page 2 (Disclosure); 6pp; Chinese.
 XX
 CC The present invention belongs to the field of human genome technology.
 CC The present invention describes the human testis mitotic centromere-
 CC associated kinesin NYD-YIF2 gene which has a cDNA sequence of 2334 base
 CC pairs (bp) containing an open reading frame sequence of 1815 bp from
 CC position 135 to 2156, encoding a sequence of 604 amino acids and has a
 CC Genbank number of AF333335. The present invention also describes:
 CC (1) utilising the NYD-KIF2 gene to prepare a fusion protein;
 CC (2) utilising the protein to immunise an animal and to prepare
 CC monoclonal and polyclonal antibodies; and (3) utilising the NYD-KIF2
 CC gene in preparing a testicular specific function gene expressing chip.
 CC The expressed protein may become gene medicine for treating related
 CC disease. The present sequence represents the human testis mitotic
 CC centromere-associated kinesin NYD-YIF2 protein from the present
 CC invention.
 XX Sequence 604 AA;
 XQ

Query Match	9.7%	Score 663	DB 23	Length 604
Best Local Similarity	44.7%	Pred	No. 1.9e-29	
Matches 149	Conservative 61	Mismatches 115	Indels 8	Gaps 6

QY	4	KIKVWVRKRP	LELEKEKKDSII	IVKNNCTI	YDEPRK	VDMTKK	IERIEF	IVDVDFD	6
Db	213	RICVCVRKRP	PLNQREFTLL	KDOLIIF	VSNNVWV	WHSKQK	VDLTRIQNGT	FCFDBAFDD	272
QY	64	TYDNFTY	VENTKPLI	IIDYENG	CVCS	CPAYGOTSG	KTYTLMG	-QPIYGQSTP	PGYFOY 122
Db	273	KASNLV	VOFQA	PLVEESI	FRKG	-MATCAV	GYTRSGKTYT	VAGDPS	GTARQDSC
QY	123	AAGDI	FTEFL	-NIYD	KNTKGI	FISFEY	EICGLYDLQ	KKRKVAAL	ENGKEV
Db	332	VAQDF	ELLRNST	YEKDLK	-VYGT	FTEYIG	KVYD	LLMWKKYLQ	VLEDNQDIIQVVG
QY	181	ILRVLT	KEBLL	KMDGV	LTKIGV	NSQND	ESSRSH	ALLNTD	KDINKNTSLG
Db	391	EKEVC	VEEVLN	LV	EGNS	CRTSRQ	TSVVA	HSRS	HAVPQIILK
QY	241	AGSER	GADV	SONKQ	QOTD	GANIN	RSLLAL	KECIR	AMDSDKNI
Db	449	AGNER	GADTT	KASRK	QLEG	AEINK	SLAL	KECTI	LALGQXKPH
QY	301	GK-SK	SIM	IANIS	PTIS	SCCE	QTLANT	RIYS	SRVK 332
Db	509	GONS	STCM	ATIS	PGMT	SCENT	INTL	TRYAN	RVRK 541

RESULT 34
AAP60452
ID AAP60452 standard; Protein, 537 AA
XX
AC AAP60452;

XX				
DT	25-MAR-2003	(updated)		
PT	27-JUN-1991	(first entry)		
XX				
DE	Sequence of the Asparagine-Rich Protein (ARP)			
DE	Aq319 of Plasmodium falciparum.			
XX				
KW	Malaria vaccine.			
XX				
OS	Plasmodium falciparum:			
XX				
PV	MO8606075-A.			
XX				
PD	23-OCT-1986.			
XX				
PF	11-APR-1986;	66WO-AUO0092.		
XX				
PR	25-JUL-1985;	85AU-0001640.		
PR	11-APR-1985;	85AU-0000108.		
PR	11-APR-1985;	85AU-0001008.		
XX	11-APR-1986;	86AU-0056037.		
PA	(HALL-) HALL INST MEDICAL RES W & E. (KEMP/) KEMP D J.			
PI	Kemp D,	Anders R,	Brown G,	Coppel RL;
DR	N-PSSB;	AAI60392.		
XX				
PT	New DNA molecules encoding Plasmodium falciparum antigens - namely the SHARP ARP and MESA antigens			
PS	Claim 16; Fig 8; 52pp; English.			
CC	The antigens of the invention are prod. using recombinant DNA techniques. They may be used in vaccine compns. to stimulate an immune response against P.falciparum, the cause of human malaria. (Updated on 25-MAR-2003 to correct PA field.)			
SQ	Sequence	537 AA;		
<hr/>				
Query Match	8.8%; Score 604.5;	DB 7;	Length 537;	
Best Local Similarity	29.1%;	Pred. No. 3.2e-26;		
Matches 209;	Conservative 80;	Mismatches 193;	Indels 235;	Gaps 40.

SQ Sequence 537 AA;

Query Match	8.8%	Score 604.5	DB 7	Length 537
Best Local Similarity	29.1%	Pred. No. 3.2e-26		
Matches	209	Conservative	80	Mismatches 193; Indels 235; Gaps 40
QY	454	HNANNNNNNNNDNNNN---NDNNNNNNNDSSMYNNNMIMINNNINNNINNNNNNN	509	
DB	1	YNNNNNNNNNDGDMGPFKNQYNNNNYKFEDEEHLR-----NNSIMMNSNENEN	49	
QY	510	NNNNSHNNHLPQPYAFPTDSDFSLDDMNCILNNNDKSFILHKKLLDNTKLKRRSSCD	569	
DB	50	ENENK-----NY-HTGSSSMKQTD-----NKKTDTSLR-----NI-----	79	
QY	570	NIMKKKKNNLHLAHSVSSKLTWESYDPOKNKQNTFFKSNINKMEDNTPKDILYESRVS	629	
DB	80	NNINNEYK-----GSAKT-----QNNND-----NNQTNEDNN-----YQNNMGN	112	
QY	630	NNMG-NNVLGLANKTTHDITKDEHNHNDKILNNGYINILNNSNVSINNSMNSINNSN-	687	
DB	113	RNNSINIMQG-----SODKQNSN--NNFYIN--QNTNPFYKQNNKNNNSNNT	157	
QY	688	---MNSNGIYSNYSNQS-----ISDQIRYVENMDSNK-----NNDNIFFDALISCDN	734	
DB	158	MHSRNSNVI--NNYTNQOTRFSSEFMNNNHHFKKKYDNNNNMKGTINDDNNMDY-----N	211	
QY	735	NNYPRITNNNN---NNNNNNNNNNIDVEYTNRRDGTNNSMKLYAYNSHLFQGDN---	788	
DB	212	KFYTYMKNKLKFNNNNNNNNNNGVFSNRMLNNNNNINNNNNKNNYNYQNRKSNMNNM	271	
QY	789	NKNTSNINONITNNKNOGQNVYSNN-----FCHYNLDKQYLYLDLNKKEQCDKIHGCD	843	
DB	272	NN	327	

QY 626 RNVSNNNGVLLGLKQTHHDISTKDEHNDKINNGVITINNSVNSINNSMNSINTN 685
DB 95 -----NNNNNNNNNN-----NNNNNNNNNNNNNNNNNNNNNNNNNNNN 121
QY 686 SNNNSNSIYKSNYSNOSISDVQIRYVENMDTSKNKNIDIFDAISCDNNMYPIITNNNN 745
DB 122 NNN 147
QY 746 NNN 805
DB 148 NNN 187
QY 806 DGNVYSNFCCHYNLDKQVLLIDLNKKEQDKNIHGCNDNIIONRDEKKKKTFFYNNN 865
DB 188 NNNNN-----NNNN-----NNN 195
QY 866 NIVIVNNNNNGNNSPRMKYGLGSHSTIDNNKNNKNNKNNKNNKNNKNNSS 925
DB 196 N-----NNNNNNNNNN-----NNNNNNNNNN-----NNNNNNNN 221
QY 926 SSSNNNNIYNNINDDTFQNDYCHNDFTIRKKNITNINSIYONDIIYITINSIDYM 985
DB 222 NNN 242
QY 986 SNTLLHFKKXYTPTLSTNEDIYNNKMEGKHRLDDQDKYDDNNNNVNNNNKNNVNNV 1045
DB 243 -----NN 260
QY 1046 DNNVVNNVNDNNKNNVNDNNVDDDDVDVFNHKNFN 1083
DB 261 NNN 298

RESULT 40
AAB18218
ID AAB18218 standard; Protein; 2010 AA.
AC AAB18218;
DT 07-NOV-2000 (first entry)
DE Plasmodium falciparum chromosome 2 related protein SEQ ID NO:75.
XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
KM antimalarial; malaria; protozoacide; infection; insecticide.
XX Plasmodium falciparum.
OS WO200025728-A2.
XX 11-MAY-2000.
PD 05-NOV-1999; 99WO-US26796.
PF 05-NOV-1998; 98US-0107131.
PR 05-NOV-1998;
PA (HOEF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
PI Hoffman S, Carucci D, Gardner M, Venter JC;
DR WPI; 2000-365347/31.
XX Proteins encoded by chromosome 2 of the human malarial parasite,
PT Plasmodium falciparum, useful as antimalarial vaccines and in the
PT diagnosis of P.falciparum infection -
XX Disclosure; Page 177-182; 577pp; English.
XX The present invention describes proteins and their fragments (I) encoded
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.

CC Also described are: (I) nucleotide sequences (II) encoding (I); and (2)
CC vaccines against P. falciparum infection comprising (I) or (II).
CC (I) and (II) are useful for the development of vaccines against
CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal
CC antibody raised to immunogens comprising the sequences of (I), are
CC useful in the detection of infection with P. falciparum. Furthermore,
CC (I) (especially when they are rifin or secreted or membrane proteins)
CC can aid the identification of drugs to treat or prevent P. falciparum
CC infection, or they can be used to identify drug resistance in
CC P. falciparum. Sequencing of the plasmodium chromosome 2 and the
CC subsequent identification of proteins encoded by it will help to expand
CC our understanding of parasite biology, a process hampered by the
CC complexity of the parasite life cycle, and provide new targets for
CC vaccine and drug development. Parasite resistance to drugs and mosquito
CC resistance to insecticides have led to a resurgence of malaria in many
CC parts of the world, and there is a pressing need for vaccines and new
CC drugs. AAB70078 to AAB70287 and AAB18144 to AAB18352 represent nucleotide
CC and protein sequences given in the present invention, but which are not
CC specifically mentioned within the specification.

SO Sequence 2010 AA;
Query Match 8.3%; Score 569.5; DB 21; Length 2010;
Best local similarity 21.3%; Pred. No. 1,4e-23;
Matches 287; Conservative 192; Mismatches 404; Indels 467; Gaps 59;

QY 224 KDINKNTSLKINAFIDLAGSRGADTVSONKOTQ---TDGANINRSLLALKECIRADSD 280
DB 8 KEVKKKEKEK-----KGAKDKGDLFLSKKKRKKKESOKIDRYL-----INSCDSN 56
QY 281 KNHP-----FRDSELTLYRLDIFVKSISIMIANISP----- 313
DB 57 KSNVSCYLNNECFVKNISICKKCMFSEFKFQVNTKYIVRHGARTPKKIKNIWPFKEG 116
QY 314 ---TISCEQTLTLTYSSRVKNFKAKSTCINEEDTNTN---RISLIDSKGSEMANSSI 367
DB 117 KGDLTFPLGPOOSIVKGEYLKRYTYTFNK---LNKKYKRRRGRLINNKKGKGYIKKNCV 173
QY 368 ENVVYKSNHLLSNNNNNKIRGKINDKI---ERNNIKNSKSPDKREGFTSGKYSLSN 424
DB 174 KCKCTLYKKNYNN 231
QY 425 D----- 425
DB 232 ELENRKRLKHPYLRFIYEEKYFLKINRSNGHQRVFIKIRRRNNILKIMIHQL 291
QY 426 IDKIKKKKKGLINYSKSTLYNDNTINK-----HNNNNNNNNNNNNNNNNNNNN 473
DB 292 INKMKKIKKNMNN 351
QY 474 NNNNDSSWMNNMIMHMINNN 522
DB 352 NNNNDNN-----NNNDNN 400
QY 523 NYAFDTSDSSSL-----DMMNCHLNNDKSLF-----LAKKRLDNITKLKRRSS 567
DB 401 SFNVAADLKYTKYYKKNILKDKKVIYNNKKKELFFPLMEHLVYKKKLLIN-KKKKK- 457
QY 568 CDNIMNKKK-----NHLARHVSQS---KIT-----MEYDPOKN 600
DB 458 --NIKKKKKKKYDKIKYLKIKYLCIKITNSERCKLTAVGIIIGLIGISEIYIYFFP 510
QY 601 KDNTFPKSNINKNMEDNTPKDIYESRVNVMNGVLLGLKQTHHDISTDQENNDKIN 660
DB 511 --ILFKNVDKTDNDNNIDITYTKKKKKCKLNKSKCFQWILNRDITSGQYNCID-KNT 567
QY 661 NGVINIINNSVNSINNSNNN-----SINNSNNSIYKSN-YNSNOSISD 706
DB 568 APVKNYIIIGENLGCENCGKNGCGDILRGDILGIDLRGNNSIPLFRSNRIFCKQS--- 624
QY 707 VOIRYVENMDTSKNKNIDIFDAI-----SCDNNMYPIITNNNNNNNNNNNNNNNN 761
DB 625 -KITFCDEL-----YIFNKIKIRLOSID-DMY-----KINHEVK 657

QY 762 NY--NNRDTNNS--MKLYAYNSHNLFOPDNNKNTSNIQININTNKQNDGNNVYGMNFCBYN 819
 Db 658 MFGNDKDVLNNSYKCYDKNDYGSY--PSYKYSNDYNS-----HYV 697
 QY 820 LNDKNTLIDLNNKQKDKNTHG--CDNN--IIONRNDPEKKKK-----TNFYNNNNIV 868
 Db 698 I-----KKKQNVKSVQCSNESIILKEROENKKEKKKKKKMENTEPIINNINLM 744
 QY 869 -----IYNNNNGNNSPRMKYGLCGSHSIDNNKNNEMKNNEMKDNHMKSN 920
 Db 745 YNINVPFDLIINERGN-----FOFYNNIKKKQKQNEK-----777
 QY 921 NNSSSSSSSNNNTYNNINDDTFO-----NDYCHNDNTFTIRKNNNTNINSNIYOND 973
 Db 778 -----GLEEMNYN-----IFQLYMKYILNESKFEKLPKFLNKVVENIDN-----818
 QY 974 IITYINSL-----NDYNSNTLLHFEKETYFTPLSTNEDIYK--EMEGKHRLDDQDKYDN 1028
 Db 819 ---TFNSITNIYKYYINMVFVHRKDCFEKQIHSKEHMKKIHRLDKFIEYEKENEIIDN 875
 QY 1029 DNN--NVQNNK---NNVDNNVDNNNVQNNVDNN-----K 1059
 Db 876 CNNINNDNKKKEIINNNTNNINNDNNIEIDSSNNFIETYYIFYLNNMTYIOFLFYELK 935
 QY 1060 NN-----VDNNVDNDDDDVDVFNINQNFNN-----EYLSYFOKNVDTII---1099
 Db 936 NTYILFSVYKVAERNISMLKTLKTKNHYIKKLNHIHNSDVYKILNNYKDELFIVYDI 995
 QY 1100 ---NCLNSLDISSMYDTK---EILNNIILSKYKAEKDNVYIKYINEDIKNSLEE 1150
 Db 996 TKWTENCMNTDI--LYNDVKNTKIDLENIIDIPITNDKEEY-----HYNNSIISV 1046
 QY 1151 IDTAQSIYERKRVLLTKLLLFKQAVDIOINNETSDLRDLYMCHICNNPPDDQEHFYA 1210
 Db 1047 LKQNSSVYKLLKKLKNSTIL-----KDLKK--LNCNFINKNYIHNNTYDK 1090
 QY 1211 YSRLEKDIINLIMRQIWCESENRLIYQFLVVEYQKNSANSVLANVSSNN-----1261
 Db 1091 HNKIYQDKIKQWYTHPNNKKQVKKIIFK--ISAYDAYIYHGVNLNFPRAVEKLSQHP 1149
 QY 1262 -GDIILINKLVODN-----IKNSMDHNN 1284
 Db 1150 PSSIDILIKKEYGQNNYIINGEIKKEYEONN 1179

Search completed: October 2, 2003, 16:17:06
 Job time : 105 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 2, 2003, 16:28:56 ; Search time 52 Seconds
(without alignments)
2382.024 Million cell updates/sec

Title: US-10-006-780-2

Perfect score: 1288
Sequence: 1 NMSKIKVVRKRPLSLEKK.....KLLVDNINSMNDNNHKK 1288

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	1.3	391	2	SI4577
2	17	1.3	419	2	SI4508
3	17	1.3	666	2	S50452
4	17	1.3	749	2	S23467
5	17	1.3	758	2	S45825
6	17	1.3	943	2	S68824
7	17	1.3	1093	2	T18275
8	17	1.3	1139	2	S71092
9	17	1.3	1178	2	S57698
10	17	1.3	1338	2	T18287
11	17	1.3	1457	2	T14577
12	17	1.3	1671	2	S71628
13	17	1.3	1969	2	T08875
14	17	1.3	4550	2	T18440
15	17	1.2	188	2	T02389
16	17	1.2	317	2	S05356
17	16	1.2	328	2	T02046
18	16	1.2	373	2	T52593
19	16	1.2	401	2	T18285
20	16	1.2	490	2	A46391
21	16	1.2	490	2	S52830
22	16	1.2	495	1	S52641
23	16	1.2	566	2	S53813
24	16	1.2	580	2	T18439
25	16	1.2	600	2	T18446
26	16	1.2	686	2	A71607
27	16	1.2	720	2	T51007
28	16	1.2	758	2	S54522
29	16	1.2	831	2	T08611

30	16	1.2	947	2	T08605
31	16	1.2	961	2	S67568
32	16	1.2	962	2	JC5808
33	16	1.2	964	2	S48404
34	16	1.2	1075	2	S54067
35	16	1.2	1225	2	T09057
36	16	1.2	1256	2	SI4556
37	16	1.2	1271	2	T08607
38	16	1.2	1336	2	T18288
39	16	1.2	1365	2	SI4871
40	16	1.2	1390	2	T14004
41	16	1.2	1402	2	T17456
42	16	1.2	1584	2	T18276
43	16	1.2	1585	2	T18274
44	16	1.2	1619	2	T18499
45	16	1.2	1743	2	T18279
46	16	1.2	1817	2	D71606
47	16	1.2	1844	2	D71612
48	16	1.2	1855	2	S41649
49	16	1.2	1858	2	T18273
50	16	1.2	1905	2	T18267
51	16	1.2	2150	2	S71629
52	16	1.2	2380	2	E71604
53	16	1.2	2523	2	T18477
54	16	1.2	3844	2	T18402
55	16	1.2	3848	2	T17414
56	16	1.2	4981	2	T18489
57	15	1.2	125	2	E43692
58	15	1.2	421	2	A23535
59	15	1.2	612	2	S65196
60	15	1.2	837	2	S36713
61	15	1.2	888	2	S64016
62	15	1.2	1655	2	T13998
63	15	1.2	2278	1	S55274
64	15	1.2	2539	2	B71619
65	15	1.2	3394	2	T18501
66	14	1.1	150	2	D48835
67	14	1.1	247	2	T34214
68	14	1.1	264	2	A44969
69	14	1.1	375	2	T00929
70	14	1.1	668	1	T02791
71	14	1.1	700	1	S09748
72	14	1.1	744	2	T06048
73	14	1.1	754	2	S48020
74	14	1.1	777	2	C85065
75	14	1.1	793	2	S34830
76	14	1.1	832	2	T36749
77	14	1.1	858	1	A42239
78	14	1.1	913	2	T18503
79	14	1.1	955	2	A47334
80	14	1.1	987	2	B96766
81	14	1.1	1056	2	C96661
82	14	1.1	1245	2	D71613
83	14	1.1	1662	1	H71402
84	14	1.1	1770	2	S56221
85	14	1.1	1891	2	T13594
86	14	1.1	2010	2	B71616
87	13	1.0	201	2	S54475
88	13	1.0	367	2	T46021
89	13	1.0	426	2	T04318
90	13	1.0	554	2	T50118
91	13	1.0	608	2	T18437
92	13	1.0	669	2	S64795
93	13	1.0	697	2	T16306
94	13	1.0	708	2	A51185
95	13	1.0	728	2	S21913
96	13	1.0	851	2	S67285
97	13	1.0	904	2	S69691
98	13	1.0	1065	2	T13230
99	13	1.0	1072	2	T13232
100	13	1.0	1074	2	T13229
101	13	1.0	1075	2	T45570
102	13	1.0	1081	2	T13231

hypothetical prote
probable membrane
G protein-coupled
probable membrane
probable membrane
asparagine-rich pr
hypothetical prote
ABC transport prot
suppressor two of
trfa protein - sli
cell surface prote
protein-tyrosine k
1-phosphatidylinos
hypothetical prote
multidrug resistan
hypothetical prote
hypothetical prote
DNA polymerase - m
1-phosphatidylinos
multidrug resistan
sensory transducti
hypothetical prote
hypothetical prote
asparagine/asparta
Tpc protein - sli
hypothetical prote
T3C protein - rabb
clustered asparagi
probable membrane
CCR4 protein - yea
probable regulator
gene mastermind pr
FAB1 protein - yea
hypothetical prote
hypothetical prote
kinesin-like prote
hypothetical prote
circumsporozoite p
squamosa-promoter
mitotic centromere
kinesin-related pr
kinesin-related pr
kinesin-related pr
kinesin-like prote
adenylate cyclase
hypothetical prote
Lckin kinesin-rela
protein kinesin-F2
kinesin-like prote
GAF domain protein
probable kinesin -
hypothetical prote
hypothetical prote
phosphatase (acid
hypothetical prote
hypothetical prote
homeobox protein L
kinesin-related pr
hypothetical prote
suppressor protein
hypothetical prote
G-box-binding fact
Brcore-TNT1-Q1-Z1
NUP1 protein - yea
hypothetical prote
dachshund isoform
dachshund protein
dachshund protein
kinesin-like prote
dachshund protein

103	13	1.0	1083	2	T26844	hypothetical prote
104	13	1.0	1151	2	A41529	GRI1 protein - yea
105	13	1.0	1335	2	T18289	racGp protein - s
106	13	1.0	1920	2	T13893	gene hindsight pro
107	13	1.0	2269	2	T18472	hypothetical prote
108	13	1.0	2573	2	D11614	hypothetical prote
109	12	0.9	78	2	S23689	erythrocyte membra
110	12	0.9	212	2	S49807	hypothetical prote
111	12	0.9	239	2	S14470	asparagine-rich pr
112	12	0.9	431	2	S50977	hypothetical prote
113	12	0.9	537	2	A23770	asparagine-rich pr
114	12	0.9	590	2	A26638	homeoic protein D
115	12	0.9	603	2	F96797	probable GT-like t
116	12	0.9	647	2	S06450	steroid hormone re
117	12	0.9	671	2	T10755	kinesin-related pr
118	12	0.9	716	1	A44259	kinesin-related pr
119	12	0.9	747	2	T23607	hypothetical prote
120	12	0.9	764	2	H71607	hypothetical prote
121	12	0.9	813	2	T46242	kinesin-like prote
122	12	0.9	914	2	S46593	finger protein AZF
123	12	0.9	1341	2	S50366	probable membrane
124	12	0.9	1516	2	E71619	RAD2 endonuclease
125	12	0.9	1570	2	T18272	1-phosphatidylinos
126	12	0.9	1864	2	T18485	hypothetical prote
127	12	0.9	2452	1	RN202L	DNA-directed RNA p
128	12	0.9	2910	2	T28156	DNA-directed RNA p
129	12	0.9	3724	2	T18437	hypothetical prote
130	11	0.9	95	2	F41298	kinesin-like prote
131	11	0.9	234	2	S14469	asparagine-rich pr
132	11	0.9	280	2	C68317	protein T10022.23
133	11	0.9	330	2	A48835	kinesin-like prote
134	11	0.9	331	2	G96785	protein F10A5.24 (
135	11	0.9	443	2	A54813	cAMP receptor CAR4
136	11	0.9	448	2	S05355	hypothetical prote
137	11	0.9	461	2	T10265	arabinogalactan-pr
138	11	0.9	472	2	A26357	homeoic protein C
139	11	0.9	519	2	I54523	kinesin-related pr
140	11	0.9	522	2	T18504	hypothetical prote
141	11	0.9	549	2	T20720	hypothetical prote
142	11	0.9	601	2	T00119	probable transcrip
143	11	0.9	622	2	A57281	kinesin-like motor
144	11	0.9	672	2	T29898	kinesin protein OS
145	11	0.9	672	2	S54351	kinesin osm-3 - Ca
146	11	0.9	676	2	S61977	transcription fact
147	11	0.9	690	2	T34149	hypothetical prote
148	11	0.9	703	2	A86319	P25116.11 protein
149	11	0.9	722	2	D71607	VP545-like protein
150	11	0.9	729	1	A34796	kinesin-related pr
151	11	0.9	784	2	T40594	probable kinesin-1
152	11	0.9	805	2	S64238	kinesin-related pr
153	11	0.9	817	2	T50240	kinesin-like prote
154	11	0.9	836	2	T18460	hypothetical prote
155	11	0.9	861	2	T00434	probable kinesin h
156	11	0.9	920	2	T08853	protein phosphatas
157	11	0.9	921	2	T01775	hypothetical prote
158	11	0.9	932	2	T30099	hypothetical prote
159	11	0.9	935	2	T51930	kinesin [imported]
160	11	0.9	968	2	T51933	kinesin motor proc
161	11	0.9	975	2	T08606	protein phosphatas
162	11	0.9	987	2	T51360	kinesin-like heavy
163	11	0.9	1025	2	T18376	multidrug resistan
164	11	0.9	1032	2	B66224	hypothetical prote
165	11	0.9	1122	2	F71606	probable multiple
166	11	0.9	1225	2	A56514	chromokinesin - ch
167	11	0.9	1226	2	I51617	kinesin-like prote
168	11	0.9	1229	2	T48959	kinesin-like prote
169	11	0.9	1230	2	T18256	probable serine/ch
170	11	0.9	1230	2	T18259	serine/threonine p
171	11	0.9	1231	2	A54803	microtubule-associ
172	11	0.9	1265	2	T03792	kinesin-related pr
173	11	0.9	1265	2	T07397	kinesin heavy chai
174	11	0.9	1269	2	T14476	psia protein - sli
175	11	0.9	1351	2	C71607	hypothetical prote
176	11	0.9	1388	2	T30335	KLP2 protein - Afr
177	11	0.9	1459	2	T30196	kinesin motor prot
178	11	0.9	1979	2	C71622	hypothetical prote
179	11	0.9	1997	2	F71607	DNA helicase II BR
180	11	0.9	2061	2	T13751	transcription fact
181	11	0.9	2339	2	A45597	DNA-directed RNA p
182	11	0.9	2457	2	T18492	hypothetical prote
183	11	0.9	2708	2	T09079	probable chloroqui
184	11	0.9	3973	2	B71612	hypothetical prote
185	10	0.8	97	2	B96580	p51.5 [imported]
186	10	0.8	100	2	S23586	erythrocyte membra
187	10	0.8	240	2	A11305	metal cations ABC
188	10	0.8	240	2	A11677	erythrocyte membra
189	10	0.8	272	2	S23684	hypothetical prote
190	10	0.8	368	2	S46601	hypothetical prote
191	10	0.8	390	1	B49070	ecdysone-inducible
192	10	0.8	398	1	R3BYM1	ribosomal protein
193	10	0.8	419	2	T18420	hypothetical prote
194	10	0.8	522	2	B45268	interleukin-9 rece
195	10	0.8	575	1	JH0827	glutamate decarbox
196	10	0.8	587	2	T19693	hypothetical prote
197	10	0.8	587	2	T29324	hypothetical prote
198	10	0.8	613	2	T15489	hypothetical prote
199	10	0.8	623	1	S33167	gene pointed prote
200	10	0.8	633	2	T04179	hypothetical prote
201	10	0.8	648	1	JQ1150	protein kinase (EC
202	10	0.8	695	2	T24950	hypothetical prote
203	10	0.8	716	2	T26998	hypothetical prote
204	10	0.8	718	1	S33168	gene pointed prote
205	10	0.8	720	2	Q01676	ARIS protein - Ara
206	10	0.8	766	2	S67044	NTII protein - yea
207	10	0.8	786	2	T18469	hypothetical prote
208	10	0.8	823	2	S48986	probable protein k
209	10	0.8	864	2	A49070	ecdysone-inducible
210	10	0.8	928	2	T16759	hypothetical prote
211	10	0.8	974	1	A40213	optic lobe develop
212	10	0.8	1068	2	S64015	pleiotropic drug r
213	10	0.8	1121	2	T13750	kinesin-like prote
214	10	0.8	1121	2	T13796	kinesin-related pr
215	10	0.8	1190	2	S47536	oxysterol-binding
216	10	0.8	1195	2	S38174	probable purine nu
217	10	0.8	1228	2	S37621	Ca2+-transporting
218	10	0.8	1254	2	T18277	kinesin heavy chai
219	10	0.8	1294	2	T18473	hypothetical prote
220	10	0.8	1308	2	E71622	probable membrane
221	10	0.8	1365	2	T18419	hypothetical prote
222	10	0.8	1419	1	DVZOF	multidrug resistan
223	10	0.8	1438	2	E71610	WD40 WEB-1 homolog
224	10	0.8	1553	2	S67483	adenosinetriphosph
225	10	0.8	1576	2	T29237	hypothetical prote
226	10	0.8	1584	1	JN0114	kinesin-related pr
227	10	0.8	1584	2	T18822	kinesin-like prote
228	10	0.8	1711	2	T18429	hypothetical prote
229	10	0.8	1737	2	A59235	unconventional myo
230	10	0.8	1817	2	H71611	probable secreted
231	10	0.8	1893	2	T22661	hypothetical prote
232	10	0.8	1921	2	T13827	kinesin-73 - fruit
233	10	0.8	2359	2	T03094	A-kinase anchor pr
234	10	0.8	2529	2	A56923	transcription fact
235	10	0.8	2578	2	A56922	regulatory protein
236	10	0.7	32	2	S08482	hypothetical prote
237	9	0.7	114	2	T17937	homeoic protein A
238	9	0.7	131	2	S33375	topoisomerase II -
239	9	0.7	160	2	S47661	hypothetical prote
240	9	0.7	172	2	T20210	FTZ-FI steroid rec
241	9	0.7	173	2	A47303	ORF MSV172 hypoch
242	9	0.7	184	2	T28333	hypothetical prote
243	9	0.7	257	2	T00406	probable integral
244	9	0.7	266	2	E71601	LMO protein - frui
245	9	0.7	266	2	S61522	hypothetical prote
246	9	0.7	300	2	H96636	hypothetical prote
247	9	0.7	305	2	T27151	probable homeodoma
248	9	0.7	321	2	G84750	

249	9	0.7	323	2	T15311	322	9	0.7	1557	2	T13160	protein CNK - fru1
250	9	0.7	344	2	T48827	323	9	0.7	1677	2	T43021	vitellogenin precu
251	9	0.7	372	2	T50014	324	9	0.7	1714	2	E71609	Ser/Thr protein ki
252	9	0.7	372	2	T29359	325	9	0.7	2206	2	G71611	hypothetical prote
253	9	0.7	375	2	A46390	326	9	0.7	2485	1	H71621	serine/threonine-s
254	9	0.7	377	1	A48711	327	9	0.7	2510	2	T28160	hypothetical prote
255	9	0.7	387	2	T25452	328	9	0.7	2657	2	T18497	hypothetical prote
256	9	0.7	388	2	D85080	329	9	0.7	2819	2	T09080	probable chloroqui
257	9	0.7	410	2	C86835	330	9	0.7	3092	2	S46009	GRPase-activating
258	9	0.7	452	2	T33049	331	9	0.7	3119	2	T18414	protein g377 - mal
259	9	0.7	472	2	S28026	332	8	0.6	58	2	S78672	hypothetical prote
260	9	0.7	508	2	E71620	333	8	0.6	65	2	T17665	hypothetical prote
261	9	0.7	508	2	A32225	334	8	0.6	115	2	D61615	sericin MG-1 - gre
262	9	0.7	518	2	S23692	335	8	0.6	117	2	A45509	kinesin-like prote
263	9	0.7	528	2	S32593	336	8	0.6	118	2	C41298	ribosomal protein
264	9	0.7	554	2	S16304	337	8	0.6	121	2	G71711	probable membrane
265	9	0.7	557	2	A71619	338	8	0.6	121	2	S67700	30S ribosomal prot
266	9	0.7	577	2	B44307	339	8	0.6	134	2	E84903	high mobility grou
267	9	0.7	578	2	S55102	340	8	0.6	145	2	H97707	regulatory protei
268	9	0.7	581	2	F84599	341	8	0.6	147	2	J01320	hypothetical prote
269	9	0.7	589	2	C85360	342	8	0.6	150	2	G64077	hypothetical prote
270	9	0.7	594	2	JC5146	343	8	0.6	151	2	T00286	hypothetical prote
271	9	0.7	598	2	T24717	344	8	0.6	163	2	T04467	hypothetical prote
272	9	0.7	602	2	B81420	345	8	0.6	182	2	T20185	hypothetical prote
273	9	0.7	640	2	A41726	346	8	0.6	186	2	T30918	hypothetical prote
274	9	0.7	642	2	T03754	347	8	0.6	188	2	G96652	protein F23N19.15
275	9	0.7	642	2	S27806	348	8	0.6	205	2	T15450	hypothetical prote
276	9	0.7	643	2	S17997	349	8	0.6	215	2	F84508	probable CCAR-box
277	9	0.7	658	2	C71856	350	8	0.6	236	2	B84718	hypothetical prote
278	9	0.7	658	2	B64659	351	8	0.6	237	2	T30920	homeobox protein -
279	9	0.7	659	2	S14535	352	8	0.6	242	2	T22915	hypothetical prote
280	9	0.7	675	2	B56262	353	8	0.6	256	2	S23691	erythrocyte membra
281	9	0.7	675	2	D81846	354	8	0.6	257	2	E89124	protein K07C11.1
282	9	0.7	675	2	F81095	355	8	0.6	259	2	G84685	probable DOF zinc
283	9	0.7	682	2	A49121	356	8	0.6	260	2	S60480	low temperature-in
284	9	0.7	682	2	A43318	357	8	0.6	265	2	B86338	protein F5M15.22
285	9	0.7	704	2	S21911	358	8	0.6	273	2	T03254	probable carbonate
286	9	0.7	708	2	T03835	359	8	0.6	274	1	JQ0957	myb-related protei
287	9	0.7	713	1	S46838	360	8	0.6	275	2	D96826	hypothetical prote
288	9	0.7	719	2	S61046	361	8	0.6	282	2	T08445	hypothetical prote
289	9	0.7	749	2	A57276	362	8	0.6	284	2	S58650	hypothetical prote
290	9	0.7	770	1	A44337	363	8	0.6	285	2	S27843	homeotic protein s
291	9	0.7	772	2	T43034	364	8	0.6	286	2	C86423	unknown protein, 1
292	9	0.7	775	2	T43033	365	8	0.6	286	2	C61615	sericin MG-2 - gre
293	9	0.7	779	1	S40382	366	8	0.6	294	2	S19226	cold-regulated pro
294	9	0.7	780	2	A48143	367	8	0.6	298	2	A32872	myogenic factor CM
295	9	0.7	784	2	S26638	368	8	0.6	307	2	A45600	asparagine-rich bl
296	9	0.7	807	2	B71605	369	8	0.6	328	1	T00747	RING-H2 finger pro
297	9	0.7	832	2	T49494	370	8	0.6	329	1	S61884	carbonate dehydrat
298	9	0.7	834	2	T06055	371	8	0.6	330	1	S61883	carbonate dehydrat
299	9	0.7	860	2	JC5702	372	8	0.6	330	1	S61882	carbonate dehydrat
300	9	0.7	868	2	JC5701	373	8	0.6	331	1	S46675	carbonate dehydrat
301	9	0.7	883	2	T18484	374	8	0.6	333	2	T51650	probable transcript
302	9	0.7	885	2	D86151	375	8	0.6	334	2	B88482	protein C05D11.6
303	9	0.7	905	2	S56295	376	8	0.6	335	2	D86410	protein F3M18.16
304	9	0.7	928	2	S50578	377	8	0.6	336	2	S04682	ribosomal protein
305	9	0.7	960	1	A39651	378	8	0.6	339	2	F47301	VirB11 homolog - B
306	9	0.7	968	2	T45746	379	8	0.6	341	2	T22551	hypothetical prote
307	9	0.7	1003	2	T13856	380	8	0.6	348	2	S19873	single-stranded DN
308	9	0.7	1070	2	S57112	381	8	0.6	354	2	A39609	glutathione trans
309	9	0.7	1091	2	S57112	382	8	0.6	356	2	G84904	probable DOF zinc
310	9	0.7	1094	2	S46021	383	8	0.6	360	2	S69063	probable membrane
311	9	0.7	1110	2	T19673	384	8	0.6	360	2	T33835	hypothetical prote
312	9	0.7	1180	2	S69205	385	8	0.6	361	2	C84613	probable cyclin D
313	9	0.7	1193	2	G71605	386	8	0.6	367	1	OZQ0MY	citromporozolite p
314	9	0.7	1221	2	T13283	387	8	0.6	369	2	S76435	hypothetical prote
315	9	0.7	1254	2	T18425	388	8	0.6	372	2	I38042	single-stranded DN
316	9	0.7	1337	2	T14594	389	8	0.6	373	2	S72368	myb-related transc
317	9	0.7	1397	2	T10466	390	8	0.6	376	2	T51673	hypothetical prote
318	9	0.7	1407	2	B42239	391	8	0.6	377	2	F84473	hypothetical prote
319	9	0.7	1436	2	D71618	392	8	0.6	383	2	S51651	hflK protein (limp
320	9	0.7	1445	2	T14913	393	8	0.6	384	2	AD3328	gene M5SP-2 protei
321	9	0.7	1553	2	T18502	394	8	0.6	389	2	S53612	

395	8	0.6	390	2	E81408	468	8	0.6	883	2	T40128	kinesin-like prote
396	8	0.6	397	2	T06531	469	8	0.6	910	2	S40259	glucose-6-phosphat
397	8	0.6	397	2	S54018	470	8	0.6	911	2	T18451	hypothetical prote
398	8	0.6	403	2	T45580	471	8	0.6	941	2	T51135	ligand-gated chann
399	8	0.6	403	2	T49480	472	8	0.6	947	2	T03795	ornithine decarbox
400	8	0.6	403	2	S47659	473	8	0.6	962	2	D86186	hypothetical prote
401	8	0.6	411	2	E96665	474	8	0.6	962	2	S03818	carboxymethylcellu
402	8	0.6	413	2	A55238	475	8	0.6	970	2	F64230	spore germination
403	8	0.6	416	2	A55237	476	8	0.6	970	2	S48485	transcription acti
404	8	0.6	429	2	S24705	477	8	0.6	976	2	T51137	ionotropic glutama
405	8	0.6	431	2	S50633	478	8	0.6	1010	2	T13167	lola-like protein
406	8	0.6	435	2	S51042	479	8	0.6	1017	2	T18488	hypothetical prote
407	8	0.6	443	2	B84449	480	8	0.6	1028	2	A56038	DNA-binding protei
408	8	0.6	446	2	T14542	481	8	0.6	1038	2	S38801	pullulanase (EC 3.
409	8	0.6	453	2	T47466	482	8	0.6	1043	2	T13733	Frz-F1 protein - f
410	8	0.6	453	2	S46033	483	8	0.6	1060	2	T30347	vitellogenin conve
411	8	0.6	461	2	JN0097	484	8	0.6	1067	2	S35423	protein kinase sgg
412	8	0.6	464	2	T45612	485	8	0.6	1099	2	T18257	phospholipase C -
413	8	0.6	471	2	T32788	486	8	0.6	1119	2	T18491	hypothetical prote
414	8	0.6	476	2	T17932	487	8	0.6	1123	2	T28139	PK4 protein kinase
415	8	0.6	479	2	T02623	488	8	0.6	1152	2	F86363	hypothetical prote
416	8	0.6	485	2	S36184	489	8	0.6	1163	2	D64315	type I restriction
417	8	0.6	486	2	F69762	490	8	0.6	1166	2	C96945	cation efflux syst
418	8	0.6	488	2	B90355	491	8	0.6	1169	2	AD1069	type I site-specif
419	8	0.6	488	2	B69825	492	8	0.6	1213	2	S16356	ovo protein - fru1
420	8	0.6	492	2	E90371	493	8	0.6	1240	2	JC5209	insulin receptor s
421	8	0.6	497	2	T06727	494	8	0.6	1242	2	J50670	hypothetical prote
422	8	0.6	507	2	S05542	495	8	0.6	1247	2	E71616	hypothetical prote
423	8	0.6	520	2	F86401	496	8	0.6	1278	2	A71609	probable secreted
424	8	0.6	532	2	T18432	497	8	0.6	1284	2	T13168	probable potassium
425	8	0.6	535	2	T17629	498	8	0.6	1314	1	TNBRV6	transcription regu
426	8	0.6	536	2	H89130	499	8	0.6	1367	2	H82874	conserved hypochet
427	8	0.6	538	2	T16034	500	8	0.6	1367	2	T18466	hypothetical prote
428	8	0.6	554	2	S37936	501	8	0.6	1384	2	S78132	DNA-directed RNA p
429	8	0.6	559	2	G82326	502	8	0.6	1398	2	H71606	hypothetical prote
430	8	0.6	566	2	S54629	503	8	0.6	1403	2	S24548	homeotic protein p
431	8	0.6	566	2	T49988	504	8	0.6	1422	2	T18404	chromatin remodeli
432	8	0.6	567	2	S58750	505	8	0.6	1436	2	S57238	forked protein 5.4
433	8	0.6	576	2	G96763	506	8	0.6	1441	2	T18422	hypothetical prote
434	8	0.6	581	2	H85214	507	8	0.6	1449	2	S57237	forked protein 5.6
435	8	0.6	583	1	S20495	508	8	0.6	1468	2	A44345	nucleoporin - rat
436	8	0.6	594	2	G98078	509	8	0.6	1475	2	S42718	nuclear pore compl
437	8	0.6	604	2	S58643	510	8	0.6	1529	2	S69688	hypothetical prote
438	8	0.6	605	2	T07123	511	8	0.6	1597	1	BVFE8L	sol protein, large
439	8	0.6	610	2	T18441	512	8	0.6	1597	2	T08428	gene small optic l
440	8	0.6	612	2	T19946	513	8	0.6	1704	2	T43141	vitellogenin I - m
441	8	0.6	612	2	T08602	514	8	0.6	1748	2	S42136	cnjB protein - Tet
442	8	0.6	628	1	T39494	515	8	0.6	1788	1	T31095	vitellogenin precu
443	8	0.6	646	1	S16654	516	8	0.6	1852	1	V0CH2	vitellogenin II pr
444	8	0.6	653	2	B81277	517	8	0.6	1965	2	T33216	hypothetical prote
445	8	0.6	655	2	A29945	518	8	0.6	2023	2	T13154	polycomb protein e
446	8	0.6	658	2	T41309	519	8	0.6	2098	2	T18397	protein CTRP - mal
447	8	0.6	669	2	D82739	520	8	0.6	2175	1	S03170	homeotic protein c
448	8	0.6	678	1	A30350	521	8	0.6	2391	2	T18410	cardamoyl-phosphat
449	8	0.6	688	2	T09954	522	8	0.6	2437	2	T18482	hypothetical prote
450	8	0.6	695	2	T39941	523	8	0.6	2500	2	G71609	hypothetical prote
451	8	0.6	721	2	C84732	524	8	0.6	2562	2	T14266	xin protein - chic
452	8	0.6	750	2	T04980	525	8	0.6	3005	2	S33642	homeotic protein z
453	8	0.6	759	2	E84538	526	8	0.6	3190	2	T13828	CREB-binding prote
454	8	0.6	763	2	S17998	527	8	0.6	3759	2	A35085	trithorax protein
455	8	0.6	769	2	T08431	528	8	0.6	43	2	AG2078	hypothetical prote
456	8	0.6	779	2	T49717	529	8	0.5	46	2	B29184	vitellogenin - duc
457	8	0.6	784	1	A55236	530	8	0.5	47	2	H81576	hypothetical prote
458	8	0.6	786	1	A47547	531	8	0.5	54	2	T10786	conserved hypochet
459	8	0.6	840	2	S77615	532	8	0.5	54	2	T06514	hypothetical prote
460	8	0.6	841	2	S69563	533	8	0.5	68	2	F69799	hypothetical prote
461	8	0.6	847	2	T12979	534	8	0.5	70	2	C90886	hypothetical prote
462	8	0.6	855	2	S47533	535	8	0.5	70	2	D85732	hypothetical prote
463	8	0.6	859	2	S64195	536	8	0.5	70	2	B64898	hypothetical prote
464	8	0.6	865	2	A25762	537	8	0.5	75	2	S37191	hypothetical prote
465	8	0.6	870	2	B81689	538	8	0.5	77	2	T41800	AcMNPV orf55 - Bom
466	8	0.6	874	2	T00491	539	8	0.5	80	2	T27603	hypothetical prote
467	8	0.6	880	2	T38083	540	8	0.5	85	2	A86282	protein F10B6.20 l

541	7	0.5	89	2	E96930	transposon related	614	7	0.5	166	2	T07089	dehydrin - soybean
542	7	0.5	97	2	S06983	hypothetical prote	615	7	0.5	166	2	F81707	hypothetical prote
543	7	0.5	98	2	S13623	ht-en protein - le	616	7	0.5	166	2	S30833	hypothetical prote
544	7	0.5	99	2	T06083	probable auxin-ind	617	7	0.5	167	2	S05545	dehydrin 3 - maize
545	7	0.5	99	2	S69516	hypothetical prote	618	7	0.5	167	2	T49638	hypothetical prote
546	7	0.5	99	2	T49636	related to cell di	619	7	0.5	167	2	AE2630	hypothetical prote
547	7	0.5	100	2	A26989	nonhistone chromos	620	7	0.5	167	2	T47385	hypothetical prote
548	7	0.5	101	2	S67034	probable membrane	621	7	0.5	168	2	A39316	hypothetical prote
549	7	0.5	102	2	T05499	auxin-induced prot	622	7	0.5	168	2	S08633	RAB-17 protein - m
550	7	0.5	102	2	S72682	hypothetical prote	623	7	0.5	168	2	C75425	hypothetical prote
551	7	0.5	102	2	G96775	hypothetical prote	624	7	0.5	169	2	F84971	hypothetical prote
552	7	0.5	102	2	S26409	protein 108 precur	625	7	0.5	169	2	T34520	hypothetical prote
553	7	0.5	104	2	S55169	hypothetical prote	626	7	0.5	170	2	AF2621	polypeptide deform
554	7	0.5	105	2	C89924	conserved hypochet	627	7	0.5	170	2	F97403	polypeptide deform
555	7	0.5	105	2	C86242	protein T16B5.1 [i	628	7	0.5	172	2	T29161	hypothetical prote
556	7	0.5	106	2	T14819	dehydrin DHN2 - so	629	7	0.5	172	2	G86463	hypothetical prote
557	7	0.5	107	2	B85356	glycine-rich prot	630	7	0.5	172	2	T36107	probable serine/ar
558	7	0.5	109	2	T12208	probable cytochrom	631	7	0.5	173	2	F70112	hypothetical prote
559	7	0.5	112	2	G72502	hypothetical prote	632	7	0.5	174	2	F85028	hypothetical prote
560	7	0.5	113	2	T48112	hypothetical prote	633	7	0.5	177	2	A84856	hypothetical prote
561	7	0.5	115	2	T07576	hypothetical prote	634	7	0.5	177	2	F96739	hypothetical prote
562	7	0.5	117	2	S48324	hypothetical prote	635	7	0.5	177	2	T37212	hypothetical prote
563	7	0.5	120	2	S65067	dehydrin DHN1 - po	636	7	0.5	178	2	T27193	hypothetical prote
564	7	0.5	121	2	E81872	hypothetical prote	637	7	0.5	178	2	T20599	hypothetical prote
565	7	0.5	123	2	S69460	hypothetical prote	638	7	0.5	179	2	B84587	probable glutaredo
566	7	0.5	124	2	S45788	probable membrane	639	7	0.5	180	2	A70170	hypothetical prote
567	7	0.5	125	2	S69871	hypothetical prote	640	7	0.5	180	2	C72676	hypothetical prote
568	7	0.5	127	1	KNMJHY	dehydrin-like prot	641	7	0.5	181	2	G96715	transcription fact
569	7	0.5	127	2	T45729	dehydrin-like prote	642	7	0.5	181	2	T46017	hypothetical prote
570	7	0.5	129	2	D41298	kinesin-like prote	643	7	0.5	182	2	G86410	protein F3M18.10 (
571	7	0.5	129	2	T00282	hypothetical prote	644	7	0.5	183	2	A56579	pre-intermoult gen
572	7	0.5	132	2	T49536	hypothetical prote	645	7	0.5	183	2	T06643	hypothetical prote
573	7	0.5	133	2	A41298	kinesin-like prote	646	7	0.5	184	2	S12095	embryonic abundant
574	7	0.5	134	2	S51087	dhnl protein - pot	647	7	0.5	185	2	H84601	probable dehydrin
575	7	0.5	135	2	T09876	dehydrin - upland	648	7	0.5	185	2	P96789	protein T23E18.12
576	7	0.5	135	2	A42098	transcription fact	649	7	0.5	185	2	T05008	hypothetical prote
577	7	0.5	139	2	S05544	dehydrin 8 - barley	650	7	0.5	187	2	S69466	hypothetical prote
578	7	0.5	139	2	T05715	dehydrin - barley	651	7	0.5	187	2	T49491	hypothetical prote
579	7	0.5	140	2	T27059	hypothetical prote	652	7	0.5	188	1	SORTAD	androgen-dependent
580	7	0.5	142	2	T34136	hypothetical prote	653	7	0.5	189	2	T02570	hypothetical prote
581	7	0.5	143	2	S05543	dehydrin 9 - barley	654	7	0.5	191	2	S29008	hypothetical prote
582	7	0.5	145	2	S04042	embryonic abundant	655	7	0.5	192	2	S59537	hypothetical prote
583	7	0.5	149	2	S19130	rab15 protein - wh	656	7	0.5	192	2	H71605	probable integral
584	7	0.5	149	2	T23179	hypothetical prote	657	7	0.5	193	2	T19388	hypothetical prote
585	7	0.5	150	2	H96756	hypothetical prote	658	7	0.5	193	2	T00585	hypothetical prote
586	7	0.5	150	2	S55882	CCH finger protei	659	7	0.5	195	2	H96532	hypothetical prote
587	7	0.5	151	2	U02282	negatively phycoc	660	7	0.5	196	2	H82910	hypothetical prote
588	7	0.5	151	2	S11848	rab1d protein - r	661	7	0.5	196	2	G46326	hypothetical prote
589	7	0.5	153	2	S67294	hypothetical prote	662	7	0.5	196	2	T46326	hypothetical prote
590	7	0.5	154	2	T14727	dehydrin DHN1 - so	663	7	0.5	197	2	S55621	hypothetical prote
591	7	0.5	154	2	A05187	hypothetical prote	664	7	0.5	197	2	T25629	hypothetical prote
592	7	0.5	154	2	S55017	hypothetical prote	665	7	0.5	197	2	C88809	protein Y43C58.1 (
593	7	0.5	155	2	S43775	desiccation-relate	666	7	0.5	197	2	A90365	hypothetical prote
594	7	0.5	156	2	F85061	hypothetical-relate	667	7	0.5	197	2	A97060	N-terminal truncat
595	7	0.5	157	2	T10489	dehydrin 1 - Comme	668	7	0.5	197	2	H85065	hypothetical prote
596	7	0.5	158	2	A75556	hypothetical prote	669	7	0.5	198	2	S07130	casein B precursor
597	7	0.5	159	2	T05279	transcription fact	670	7	0.5	198	2	A49243	GLOF-C4 protein -
598	7	0.5	159	2	T48956	hypothetical prote	671	7	0.5	199	2	H84015	hypothetical prote
599	7	0.5	160	2	B86614	CT832 hypothetical	672	7	0.5	199	2	T14468	pollen coat protei
600	7	0.5	160	2	T00836	probable dnau prot	673	7	0.5	200	2	SC4357	HMGI protein - sea
601	7	0.5	160	2	F72010	CT832 hypothetical	674	7	0.5	200	2	S64905	hypothetical prote
602	7	0.5	161	2	S05547	dehydrin 17 - barl	675	7	0.5	200	2	T44143	DR3 protein (impor
603	7	0.5	161	2	T45055	hypothetical prote	676	7	0.5	200	2	T43953	hypothetical prote
604	7	0.5	161	2	D86418	unknown protein, 6	677	7	0.5	200	2	T09783	dehydration-induct
605	7	0.5	162	2	C85356	glycine-rich prote	678	7	0.5	202	2	D85058	hypothetical prote
606	7	0.5	163	2	S01152	RAB21 protein - ri	679	7	0.5	202	2	T11744	dehydrin - kidney
607	7	0.5	163	2	B81530	N utilization subs	680	7	0.5	203	2	T39117	hypothetical prote
608	7	0.5	164	2	T15525	hypothetical prote	681	7	0.5	203	2	S60449	probable membrane
609	7	0.5	164	2	S11846	rab16B protein - r	682	7	0.5	204	2	B71433	hypothetical prote
610	7	0.5	164	2	S11847	rab16C protein - r	683	7	0.5	205	2	T05713	hypothetical prote
611	7	0.5	164	2	T14187	hypothetical prote	684	7	0.5	206	2	G85065	hypothetical prote
612	7	0.5	164	2	T16168	hypothetical prote	685	7	0.5	206	2	T23852	hypothetical prote
613	7	0.5	165	2	T03392	probable dehydrin	686	7	0.5	206	2	T34347	hypothetical prote

687	7	0.5	207	2	S58744	NADH2 dehydrogenas	760	7	0.5	250	2	S36769	ubiquitin-protein
688	7	0.5	207	2	S55200	hypothetical prote	761	7	0.5	251	1	B64456	bioc protein homol
689	7	0.5	207	2	B80140	hypothetical prote	762	7	0.5	252	2	G84450	hypothetical prote
690	7	0.5	209	2	E86517	hypothetical prote	763	7	0.5	252	2	H84423	hypothetical prote
691	7	0.5	209	2	T07779	dehydrin homolog C	764	7	0.5	252	2	A83276	hypothetical cytos
692	7	0.5	210	2	B96694	hypothetical prote	765	7	0.5	254	2	S50834	emerin - human
693	7	0.5	210	2	T49785	hypothetical prote	766	7	0.5	255	1	WJHUV4	homeotic protein H
694	7	0.5	211	2	D96507	hypothetical prote	767	7	0.5	255	2	AB1942	hypothetical prote
695	7	0.5	211	2	T18799	hypothetical prote	768	7	0.5	255	2	T05958	cold acclimation p
696	7	0.5	211	2	E85091	hypothetical prote	769	7	0.5	257	2	G84712	hypothetical prote
697	7	0.5	212	2	T48580	hypothetical prote	770	7	0.5	258	1	WMV235	3K major secreted
698	7	0.5	213	2	S35522	heterochromatin pr	771	7	0.5	258	2	S71561	drought-induced pr
699	7	0.5	214	2	F84853	hypothetical prote	772	7	0.5	258	2	T05194	hypothetical prote
700	7	0.5	216	2	T29039	hypothetical prote	773	7	0.5	259	2	A84476	hypothetical prote
701	7	0.5	216	2	B96794	unknown protein FI	774	7	0.5	259	2	S55884	CCHH finger protei
702	7	0.5	216	2	S50766	dehydrin-related p	775	7	0.5	259	2	T47627	hypothetical prote
703	7	0.5	216	2	S05575	sporozite antigen	776	7	0.5	259	2	AB1080	conserved hypotet
704	7	0.5	217	2	T06455	Myb26 protein - ga	777	7	0.5	259	2	S64423	probable membrane
705	7	0.5	218	2	B86750	hypothetical prote	778	7	0.5	260	2	B64563	hypothetical prote
706	7	0.5	218	2	JC7220	nuclear protein SR	779	7	0.5	260	2	D86266	probable C2H2-type
707	7	0.5	220	2	S12849	coat protein TPa -	780	7	0.5	260	2	B96686	myb-related protei
708	7	0.5	221	2	T04271	probable chlorodox	781	7	0.5	261	2	T06650	thiol-disulfide in
709	7	0.5	222	2	G86168	hypothetical prote	782	7	0.5	261	2	E85439	hypothetical prote
710	7	0.5	224	2	S78129	ribosomal protein	783	7	0.5	261	2	T09075	hypothetical prote
711	7	0.5	224	2	H72346	basal-body rod mod	784	7	0.5	262	2	T40941	hypothetical serin
712	7	0.5	225	2	S05546	dehydrin 18 - harl	785	7	0.5	262	2	T02203	finger protein Dof
713	7	0.5	225	2	B84653	TINY-like AP2 doma	786	7	0.5	262	2	T14807	hypothetical prote
714	7	0.5	225	2	S59852	DNA-binding protei	787	7	0.5	262	2	S43953	cold acclimation p
715	7	0.5	228	2	S19132	rab25 protein - ri	788	7	0.5	263	2	H84918	glutathione S-tran
716	7	0.5	229	2	JQ0928	coat protein - tom	789	7	0.5	263	2	S05433	corticosterone / 11
717	7	0.5	229	2	F84685	hypothetical prote	790	7	0.5	263	2	JC5271	oxygen-evolving co
718	7	0.5	229	2	T04764	hypothetical prote	791	7	0.5	263	2	H84547	hypothetical prote
719	7	0.5	229	2	T52364	hypothetical prote	792	7	0.5	264	2	T09377	hypothetical prote
720	7	0.5	229	2	UC7219	nuclear protein SR	793	7	0.5	264	2	T10637	hypothetical prote
721	7	0.5	230	2	D86352	protein T26F17.14	794	7	0.5	265	2	S48360	hypothetical prote
722	7	0.5	230	2	S71747	DAG protein precir	795	7	0.5	265	2	D96667	hypothetical prote
723	7	0.5	230	2	T39598	hypothetical serin	796	7	0.5	265	2	F85014	hypothetical prote
724	7	0.5	230	2	T26896	hypothetical prote	797	7	0.5	265	2	T29973	hypothetical prote
725	7	0.5	231	2	S17367	radi58 protein - w	798	7	0.5	266	2	T05123	cold acclimation p
726	7	0.5	231	2	A10091	probable RNA polym	799	7	0.5	268	2	T06802	hypothetical prote
727	7	0.5	233	2	T00966	hypothetical prote	800	7	0.5	269	2	C84707	hypothetical prote
728	7	0.5	233	2	H83753	hypothetical prote	801	7	0.5	270	2	F86177	protein F19P19.1 (
729	7	0.5	233	2	H97094	hypothetical prote	802	7	0.5	270	2	T22151	hypothetical prote
730	7	0.5	234	2	T49448	hypothetical prote	803	7	0.5	271	2	T13013	hypothetical prote
731	7	0.5	235	2	S48924	hypothetical prote	804	7	0.5	273	2	B64446	formylmethanofuran
732	7	0.5	235	2	D84602	hypothetical prote	805	7	0.5	273	2	F71645	hypothetical prote
733	7	0.5	235	2	T30656	hypothetical prote	806	7	0.5	274	2	S26953	photoystem II 22K
734	7	0.5	238	2	S77699	inner cell wall ma	807	7	0.5	274	2	T07393	myb-related transc
735	7	0.5	238	2	S66358	DNA-binding protei	808	7	0.5	275	2	T05595	hypothetical prote
736	7	0.5	238	2	T52505	hypothetical prote	809	7	0.5	276	2	A31864	vitellogenin III -
737	7	0.5	238	2	T32889	hypothetical prote	810	7	0.5	276	2	H97861	hypothetical prote
738	7	0.5	238	2	A97336	Na+ ABC transporte	811	7	0.5	276	2	F95259	conserved hypotet
739	7	0.5	240	2	D84645	hypothetical prote	812	7	0.5	276	2	A99525	conserved hypotet
740	7	0.5	242	2	D86343	protein T2211.13	813	7	0.5	277	2	T52513	hypothetical prote
741	7	0.5	243	2	T51989	ethylene responsiv	814	7	0.5	277	2	T19378	hypothetical prote
742	7	0.5	243	2	D84792	probable protein t	815	7	0.5	278	2	T51641	myb-related transc
743	7	0.5	243	2	D84629	hypothetical prote	816	7	0.5	279	2	T01924	hypothetical prote
744	7	0.5	244	2	A42529	B2XR protein - vac	817	7	0.5	279	2	A86312	mevalonate kinase
745	7	0.5	244	2	A96806	hypothetical prote	818	7	0.5	279	2	H89827	30K immediate-earl
746	7	0.5	244	2	T51481	hypothetical prote	819	7	0.5	280	1	EDB83T	hypothetical prote
747	7	0.5	244	2	B86197	hypothetical prote	820	7	0.5	281	2	T16813	hypothetical prote
748	7	0.5	244	2	E84885	hypothetical prote	821	7	0.5	281	2	T01612	hypothetical prote
749	7	0.5	244	2	S70176	year protein - Yer	822	7	0.5	282	2	C96687	hypothetical prote
750	7	0.5	244	2	T04700	hypothetical prote	823	7	0.5	282	2	S61663	probable membrane
751	7	0.5	244	2	D96937	SAM-dependent meth	824	7	0.5	282	2	T21222	hypothetical prote
752	7	0.5	245	2	AB0300	quorum-sensing tra	825	7	0.5	283	2	F86450	hypothetical prote
753	7	0.5	245	2	B86407	hypothetical prote	826	7	0.5	284	1	KART	alpha-casein precu
754	7	0.5	246	2	S58497	early auxin-induci	827	7	0.5	284	2	A35419	neutrophil protein
755	7	0.5	247	2	JB0307	membrane protein -	828	7	0.5	285	2	T06637	hypothetical prote
756	7	0.5	247	2	T47883	secretory protein -	829	7	0.5	285	2	G70100	hypothetical prote
757	7	0.5	249	2	E82917	methionine amlope	830	7	0.5	285	2	T18689	hypothetical prote
758	7	0.5	250	2	C71611	probable secreted	831	7	0.5	286	2	H90018	hypothetical prote
759	7	0.5	250	2	T10663	hypothetical prote	832	7	0.5	286	2	C84857	hypothetical prote

833	7	0.5	286	2	T48572	hypotheical prote
834	7	0.5	286	2	T52373	homeobox protein T
835	7	0.5	288	2	B40722	homeotic protein 1
836	7	0.5	289	2	T03395	probable lipase -
837	7	0.5	289	2	T52354	hypotheical prote
838	7	0.5	290	2	T21668	hypotheical prote
839	7	0.5	291	2	T08189	hypotheical prote
840	7	0.5	292	2	T51171	transcription fact
841	7	0.5	292	2	T48327	hypotheical prote
842	7	0.5	292	2	S41012	hypotheical prote
843	7	0.5	294	2	T05381	hypotheical prote
844	7	0.5	294	2	S38983	kinesin-related pr
845	7	0.5	294	2	A96707	transcription fact
846	7	0.5	295	2	T08408	hypotheical prote
847	7	0.5	295	2	B86371	hypotheical prote
848	7	0.5	295	2	T00399	probable AP2 domai
849	7	0.5	295	2	T23989	hypotheical prote
850	7	0.5	297	2	B84731	hypotheical prote
851	7	0.5	298	2	A86840	hypotheical prote
852	7	0.5	299	2	F84810	hypotheical prote
853	7	0.5	300	2	C84853	hypotheical prote
854	7	0.5	302	2	T01942	hypotheical prote
855	7	0.5	302	2	T01540	hypotheical prote
856	7	0.5	302	2	C86480	33.2K hypotheical
857	7	0.5	302	2	H86391	hypotheical prote
858	7	0.5	303	2	T15694	hypotheical prote
859	7	0.5	303	2	B84744	hypotheical prote
860	7	0.5	304	2	T04681	hypotheical prote
861	7	0.5	304	2	T09357	hypotheical prote
862	7	0.5	304	2	T16535	hypotheical prote
863	7	0.5	305	2	T47844	hypotheical prote
864	7	0.5	305	2	JC4525	nucleic acid-bindi
865	7	0.5	306	2	S33538	embryogenic cell p
866	7	0.5	306	2	T47832	hypotheical prote
867	7	0.5	308	2	C84768	hypotheical prote
868	7	0.5	310	2	JC4343	uridine phosphoryl
869	7	0.5	310	2	D86423	hypotheical prote
870	7	0.5	311	2	T23846	hypotheical prote
871	7	0.5	313	2	S55899	transcription elon
872	7	0.5	313	2	S59448	hypotheical prote
873	7	0.5	313	2	S01303	homeotic protein a
874	7	0.5	313	2	T05256	hypotheical prote
875	7	0.5	314	2	T08933	hypotheical prote
876	7	0.5	314	2	A41544	developmental prot
877	7	0.5	314	2	A36195	developmental prot
878	7	0.5	314	2	T48514	hypotheical prote
879	7	0.5	315	2	JC7572	somite Mat1 protei
880	7	0.5	316	2	E96770	hypotheical prote
881	7	0.5	317	2	T39869	probable lysophosp
882	7	0.5	318	2	T32340	hypotheical prote
883	7	0.5	318	2	T20063	hypotheical prote
884	7	0.5	319	2	T43040	hypotheical prote
885	7	0.5	319	2	T23666	hypotheical prote
886	7	0.5	319	2	B86433	hypotheical prote
887	7	0.5	320	2	D86259	protein T1C24.6 l
888	7	0.5	321	2	A84792	hypotheical prote
889	7	0.5	321	2	S55640	hypotheical prote
890	7	0.5	323	2	T51751	homeotic protein o
891	7	0.5	323	2	G96544	hypotheical prote
892	7	0.5	324	2	F86192	hypotheical prote
893	7	0.5	324	2	B69521	hypotheical prote
894	7	0.5	324	2	C86426	hypotheical prote
895	7	0.5	325	2	G96718	unknown protein, 5
896	7	0.5	325	2	S46760	hypotheical prote
897	7	0.5	326	2	C96735	hypotheical prote
898	7	0.5	327	2	B84781	unknown protein f2
899	7	0.5	328	1	S10200	carbonate dehydrat
900	7	0.5	328	2	G02469	homeotic protein D
901	7	0.5	329	2	D64345	peripheral membran
902	7	0.5	329	2	T13016	hypotheical prote
903	7	0.5	332	2	C48835	kinesin-like prote
904	7	0.5	332	2	T48083	hypotheical prote
905	7	0.5	333	1	JW0097	bicoid-related hom
906	7	0.5	335	2	B84024	phosphate ABC tran
907	7	0.5	336	1	S28412	carbonate dehydrat
908	7	0.5	338	2	D72327	heat shock operon
909	7	0.5	339	2	T04900	hypotheical prote
910	7	0.5	339	2	D90106	hypotheical prote
911	7	0.5	339	2	H86210	hypotheical prote
912	7	0.5	339	2	T28178	hypotheical prote
913	7	0.5	342	2	AF3357	anthranilate phosph
914	7	0.5	343	2	G63654	cell wall lytic ac
915	7	0.5	343	2	T01374	hypotheical prote
916	7	0.5	343	2	T13995	gag protein - frr1
917	7	0.5	344	2	T05064	hypotheical prote
918	7	0.5	344	2	T40167	hypotheical prote
919	7	0.5	345	2	T06098	hypotheical prote
920	7	0.5	345	2	T07777	probable homeobox
921	7	0.5	347	2	T07108	glucan endo-1,3-be
922	7	0.5	347	2	H96504	hypotheical prote
923	7	0.5	347	2	H84608	probable RNA-bind
924	7	0.5	348	2	T28310	hypotheical prote
925	7	0.5	349	2	T41394	hypotheical serin
926	7	0.5	349	2	T29931	hypotheical prote
927	7	0.5	350	2	T07061	L-lactate dehydrog
928	7	0.5	350	2	T06689	heat shock protein
929	7	0.5	352	2	S22464	stricteosidine synt
930	7	0.5	353	2	AB2396	hypotheical prote
931	7	0.5	354	2	G71403	hypotheical prote
932	7	0.5	354	2	T49806	hypotheical prote
933	7	0.5	355	2	T04317	homeobox protein L
934	7	0.5	355	2	T38483	hypotheical prote
935	7	0.5	355	2	T24010	hypotheical prote
936	7	0.5	357	2	A54512	serine-repeat anti
937	7	0.5	358	2	C86291	hypotheical prote
938	7	0.5	360	2	T45649	hypotheical prote
939	7	0.5	362	2	T54464	RING-H2 zinc fing
940	7	0.5	363	2	T49209	leucoanthocyanidin
941	7	0.5	363	2	S60684	adenylate cyclase
942	7	0.5	363	2	A96573	protein Fl2M16.15
943	7	0.5	364	2	A84890	probable flavonol
944	7	0.5	364	2	T48188	gene NK6.1 protei
945	7	0.5	364	2	A95730	hypotheical prote
946	7	0.5	364	2	H86384	40.1K hypotheical
947	7	0.5	365	2	C72308	conserved hypothe
948	7	0.5	365	2	F85018	hypotheical prote
949	7	0.5	366	2	T10452	farneesyltransfera
950	7	0.5	366	2	T00439	hypotheical prote
951	7	0.5	366	2	A86392	hypotheical prote
952	7	0.5	367	2	E64399	hypotheical prote
953	7	0.5	367	2	S28604	Slizp protein - Ar
954	7	0.5	368	2	F69531	ATP-dependent RNA
955	7	0.5	368	2	A54430	hypoxic function t
956	7	0.5	368	2	T51200	hypotheical prote
957	7	0.5	369	2	T38659	melanoma antigen M
958	7	0.5	370	2	S59539	heat shock transcr
959	7	0.5	371	2	T06382	knox protein 1 - g
960	7	0.5	372	2	E85170	hypotheical prote
961	7	0.5	372	2	S20056	para-hydroxybenzo
962	7	0.5	375	2	T06436	probable ornithine
963	7	0.5	375	2	E96567	hypotheical prote
964	7	0.5	376	2	C85155	RING-H2 finger pro
965	7	0.5	376	2	S45763	hypotheical prote
966	7	0.5	377	1	W2WL13	E2 protein - human
967	7	0.5	377	1	MEMZP2	matrix protein - p
968	7	0.5	377	2	C64404	hypotheical prote
969	7	0.5	379	2	S42543	hypotheical prote
970	7	0.5	380	2	F89811	cystathionine gamm
971	7	0.5	381	2	T48623	hypotheical prote
972	7	0.5	381	2	C96657	hypotheical prote
973	7	0.5	381	2	C96657	hypotheical prote
974	7	0.5	381	2	T09640	protein phosphatas
975	7	0.5	382	2	F75179	DNA topoisomerase
976	7	0.5	382	2	D84766	hypotheical prote
977	7	0.5	383	2	T38443	hypotheical prote
978	7	0.5	383	2	T38442	hypotheical prote


```
979 7 0.5 383 2 T06753 zinc finger protei
980 7 0.5 383 2 T46707 proteophosphoglyc
981 7 0.5 383 2 S67462 hypothetical prote
982 7 0.5 383 2 S61195 hypothetical prote
983 7 0.5 383 2 A84552 probable thioredox
984 7 0.5 383 2 A82295 hypothetical prote
985 7 0.5 384 2 G96947 cytochrome c
986 7 0.5 386 2 S58755 ribosomal protein
987 7 0.5 386 2 F90112 DEAD box protein (
988 7 0.5 386 2 T48158 hypothetical prote
989 7 0.5 386 2 A82284 conserved hypotet
990 7 0.5 387 2 C96720 hypothetical prote
991 7 0.5 389 2 T04816 myb-related protei
992 7 0.5 391 2 S18666 KIN17 protein - mo
993 7 0.5 392 2 S11998 finger protein odd
994 7 0.5 392 2 T43490 hypothetical prote
995 7 0.5 393 2 A45017 transcription fact
996 7 0.5 393 2 T00647 glycosyl transfera
997 7 0.5 393 2 B84425 probable protein t
998 7 0.5 393 2 T48907 Stt2p protein (imp
999 7 0.5 394 2 C84905 probable extensin
1000 7 0.5 395 2 T00516 hypothetical prote
```

ALIGNMENTS

RESULT 1

S14577

asparagine-rich protein (clone 14C1) - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jun-2000

C:Accession: S14577

R:Schreiber, L.; Deutsche, U.; Storck, T.; Mueller-Hill, D.

submitted to the EMBL Data Library, December 1989

A:Reference number: S14469

A:Accession: S14577

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-391 <SCH>

A:Cross-references: EMBL:X17483; NID:G9842; PID:G9843

C:Superfamily: Saccharomyces cerevisiae ribosomal protein var1

Query Match 1.3% Score 17; DB 2; Length 391;

Best Local Similarity 100.0%; Pred. No. 8.2e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 741 TNNNNNNNNNNNNNNNNNNNN 757

DB 162 TNNNNNNNNNNNNNNNNNN 178

RESULT 2

S14508

asparagine-rich protein (clone 52C11) - Plasmodium falciparum (fragment)

C:Species: Plasmodium falciparum

C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Sep-1997

C:Accession: S14508

R:Schreiber, L.; Deutsche, U.; Storck, T.; Mueller-Hill, D.

submitted to the EMBL Data Library, December 1989

A:Reference number: S14469

A:Accession: S14508

A:Molecule type: mRNA

A:Residues: 1-419 <SCH>

A:Cross-references: EMBL:X17489; NID:G9854; PID:G9855

Query Match 1.3% Score 17; DB 2; Length 419;

Best Local Similarity 100.0%; Pred. No. 8.7e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 741 TNNNNNNNNNNNNNNNNNNNN 757

DB 25 TNNNNNNNNNNNNNNNNNNNN 41

RESULT 3

S50452

hypothetical protein YEL007w - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C>Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 19-Apr-2002

C:Accession: S50452

R:Dieterich, F.S.

submitted to the EMBL Data Library, December 1994

A:Description: Saccharomyces cerevisiae chromosome V cosmid 9871, 8199, 9867, 9495 and

A:Reference number: S50428

A:Accession: S50452

A:Molecule type: DNA

A:Residues: 1-666 <DIE>

A:Cross-references: EMBL:U18530; NID:G602367; PID:G602374; GSPDB:GN00005; MIPS:YEL007w

C:Genetics:

A:Gene: SGD:TOS9; MIPS:YEL007w

A:Cross-references: SGD:S0000733

A:Map position: 5L

Query Match 1.3% Score 17; DB 2; Length 666;

Best Local Similarity 100.0%; Pred. No. 1.3e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 742 NNNNNNNNNNNNNNNNNNNI 758

DB 351 NNNNNNNNNNNNNNNNNNNI 367

RESULT 4

S23467

probable long-chain-fatty-acid-CoA ligase (BC 6.2.1.3) - malaria parasite (Plasmodium f

C:Species: Plasmodium falciparum

C>Date: 22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 03-Jun-2002

C:Accession: S23467

R:Zhao, Y.; Kappes, B.; Yang, J.; Franklin, R.M.

Eur. J. Biochem. 207, 305-313, 1992

A:Title: Molecular cloning, stage-specific expression and cellular distribution of a pu

A:Reference number: S23466; MUID:92331669; PMID:11378403

A:Accession: S23467

A:Status: preliminary; translation not shown

A:Molecule type: mRNA

A:Residues: 1-749 <ZHA>

A:Cross-references: EMBL:X63648; NID:G9937; PID:G9939

C:Superfamily: human long-chain-fatty-acid-CoA ligase; acetate-CoA ligase homology

C:Keywords: acid-chol ligase; coenzyme A

F.103-732/Domain: acetate-CoA ligase homology <ACL>

Query Match 1.3% Score 17; DB 2; Length 749;

Best Local Similarity 100.0%; Pred. No. 1.4e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 742 NNNNNNNNNNNNNNNNNNNI 758

DB 271 NNNNNNNNNNNNNNNNNNNI 287

RESULT 5

S45825

cell division control protein CDC27 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YBL0718; protein YBL084C

C:Species: Saccharomyces cerevisiae

C>Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 11-Jan-2000

C:Accession: S45825; S45422; S28742; S59219

R:Domdey, H.; Gassenhuber, H.; Obermaier, B.; Piravandi, E.

submitted to the Protein Sequence Database, August 1994

A:Reference number: S45816

A:Accession: S45825

A:Molecule type: DNA

A:Residues: 1-758 <DOM>

A:Cross-references: EMBL:Z35845; NID:G536135; PID:NCAA84905.1; PID:G536136; MIPS:YBL084

A:Experimental source: strain S288C

R;Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.
submitted to the EMBL Data Library, May 1994
A:Description: Sequence analysis of a 78,6 kb segment of the left end of Saccharomyces ce
A:Reference number: S45387
A:Accession: S45422
A:Molecule type: DNA
A:Residues: 141-758 <OEB>
A:Cross-references: EMBL:X79489; NID:G496661; PIDN:CAA56022.1; PID:G496695
A:Experimental source: strain S288C
R;Sikorski, R.S.; Michaud, W.A.; Mooton, J.C.; Boguski, M.S.; Connolly, C.; Hieter, P.
Cold Spring Harb. Symp. Quant. Biol. 56, 663-673, 1991
A:Title: YPR proteins as essential components of the yeast cell cycle.
A:Reference number: S28742; MUID:92306611; PMID:1819514
A:Accession: S28742
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 607-619 <SIR>
R;Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.
Yeast 11, 1103-1112, 1995
A:Title: Sequence analysis of a 78.6 kb segment of the left end of Saccharomyces cerevis
A:Reference number: S59184; MUID:96076635; PMID:7502586
A:Accession: S59219
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 442-691 <OBW>
A:Cross-references: EMBL:X79489
C:Genetics:
A:Gene: SGD: CDC27; SNB1
A:Cross-references: SGD:S0000180; MIPS:YBL084C
A:Map position: 2L
C:Superfamily: unassigned tetratricopeptide repeat proteins; tetratricopeptide repeat hc
C:Keywords: nucleus
F:472-505/Domain: tetratricopeptide repeat homology <TT1>
F:506-539/Domain: tetratricopeptide repeat homology <TT2>
F:540-573/Domain: tetratricopeptide repeat homology <TT3>
F:574-607/Domain: tetratricopeptide repeat homology <TT4>
F:608-641/Domain: tetratricopeptide repeat homology <TT5>
F:642-675/Domain: tetratricopeptide repeat homology <TT6>
F:676-709/Domain: tetratricopeptide repeat homology <TT7>
F:710-743/Domain: tetratricopeptide repeat homology <TT8>

Query Match 1.3%; Score 17; DB 2; Length 758;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNNI 758
DB 373 NNNNNNNNNNNNNNNNNI 389

RESULT 6
rnpB protein, cytosolic - slime mold (Dictyostelium discoideum)
N:Alternate names: RING finger protein
C:Species: Dictyostelium discoideum
C>Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 24-Apr-1998
C:Accession: S68824
R;Kawata, T.; Steel, J.B.; Williams, J.G.
FEBS Lett. 386, 103-109, 1996
A:Title: RINGB: a Dictyostelium RING finger protein that is specifically located in matur
A:Reference number: S68824; MUID:96228044; PMID:8647262
A:Accession: S68824
A:Molecule type: mRNA
A:Residues: 1943 <KAW>
A:Experimental source: maturing cells
C:Genetics:
A:Gene: rnpB
C:Superfamily: RING finger homology
C:Keywords: zinc
F:893-937/Domain: RING finger homology <RNG>

Query Match 1.3%; Score 17; DB 2; Length 943;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 741 TNNNNNNNNNNNNNNNNN 757
DB 433 TNNNNNNNNNNNNNNNNN 449

RESULT 7
T18275
1-phosphatidylinositol 3-kinase (BC 2.7.1.137) 4 - slime mold (Dictyostelium discoideum
C:Species: Dictyostelium discoideum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18275
R;Zhou, K.; Takegawa, K.; Emr, S.D.; Firtel, R.A.
Mol. Cell. Biol. 15, 5645-5656, 1995
A:Title: A phosphatidylinositol (PI) kinase gene family in Dictyostelium discoideum. Bi
A:Reference number: Z06411
A:Accession: T18275
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1093 <ZHO>
A:Cross-references: EMBL:U23479; NID:G733527; PID:G733528; PIDN:AAA85725.1.
C:Genetics:
A:Note: PIK4
C:Keywords: phosphotransferase

Query Match 1.3%; Score 17; DB 2; Length 1093;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNNI 758
DB 282 NNNNNNNNNNNNNNNNNI 298

RESULT 8
UV-damaged DNA binding protein repp - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C>Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 01-Dec-2000
C:Accession: S71092; S21443
R;Alexander, H.; Lee, S.K.; Yu, S.L.; Alexander, S.
Nucleic Acids Res. 24, 2295-2301, 1996
A:Title: repp-The Dictyostelium homolog of the human xeroderma pigmentosum group E gene
A:Reference number: S71092; MUID:96279729; PMID:8710499
A:Accession: S71092
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1139 <ALE>
A:Cross-references: EMBL:U50042; NID:G1399511; PIDN:AA47162.1; PID:G1399512
R;Sydow, L.; Alexander, H.; Alexander, S.
submitted to the EMBL Data Library, April 1992
A:Reference number: S21443
A:Accession: S21443
A:Molecule type: mRNA
A:Residues: 860-1139 <SYD>
A:Cross-references: EMBL:X65937
C:Genetics:
A:Gene: repp
A:Introns: 654/2
C:Superfamily: UV-damaged DNA-binding protein

Query Match 1.3%; Score 17; DB 2; Length 1139;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNNI 758
DB 115 NNNNNNNNNNNNNNNNNI 131

RESULT 9
S57698

regulatory protein PHO81 - Yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein 1178; protein G8567; protein YGR233C
C:Species: *Saccharomyces cerevisiae*
C>Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 23-Mar-2001
C:Accession: S57698; S20136; S41074; J01032; S64557; S63914
R:Van der Aart, O.J.M.; Kleene, K.; Steensma, H.Y.
A:Submitted to the EMBL Data Library, June 1995
A:Description: Sequence analysis of the 43 KB CML-YLM9-PET54-SMT1-PHO81-YHB4-PFK1 region
A:Reference number: S57680
A:Accession: S57698
A:Molecule type: DNA
A:Residues: 1-1178 <VAN>
A:Cross-references: EMBL:X87941; NID:G886908; PID:G886927
R:Coche, T.; Prozzi, D.; Legrain, M.; Hilger, F.; Vandenhoute, J.
Nucleic Acids Res. 18, 2176, 1990
A:Title: Nucleotide sequence of the PHO81 gene involved in the regulation of the repress
A:Reference number: S20136; MUID:90245670; PMID:2186378
A:Accession: S20136
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-918, 920-1178 <COC>
A:Cross-references: EMBL:X52482; NID:G4139; PID:CAA36726.1; PID:G4140
A:Experimental source: strain GRF 88
R:Creasy, C.L.; Madden, S.L.; Bergman, L.W.
Nucleic Acids Res. 21, 1975-1982, 1993
A:Title: Molecular analysis of the PHO81 gene of *Saccharomyces cerevisiae*.
A:Reference number: S41074; MUID:93261834; PMID:8493108
A:Accession: S41074
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-110 <CRE>
A:Cross-references: GB:S61041; NID:G385415; PID:NAD13922.1; PID:G4261622
R:Wu, U.S.; Ao, S.Z.
Acta Biochim. Biophys. Sin. 26, 283-287, 1994
A:Title: Conformation prediction and function analysis of yeast PHO81 protein.
A:Reference number: J01032
A:Accession: J01032
A:Molecule type: DNA
A:Residues: 1-39, 41-78, 80-119, 121-159, 161-199, 201-230, 232-279, 281-282, 'U', 284-319, 321-357, 881-892, 'W', 894-919, 922-960, 962-1000, 1002-1040, 1042-1080, 1082-1103, 'V', 1105-1120, 1122
R:Van der Aart, O.J.M.; Steensma, H.Y.
Submitted to the Protein Sequence Database, May 1996
A:Reference number: S64541
A:Accession: S64557
A:Molecule type: DNA
A:Residues: 1-1178 <VAM>
A:Cross-references: EMBL:Z73018; NID:G1323420; PID:G1333421; MIPS:YGR233C
A:Experimental source: strain S28C
R:Van der Aart, O.J.M.; Kleene, K.; Steensma, H.Y.
Yeast 12, 385-390, 1996
A:Title: Sequence analysis of the 43 kb CML-YLM9-PET54-DIE2-SMT1-PHO81-YHB4-PFK1 region
A:Reference number: S63914; MUID:96267763; PMID:8701610
A:Accession: S63914
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1178 <VAF>
A:Cross-references: EMBL:X87941; NID:G886908; PID:CAA61183.1; PID:G886927
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995
C:Genetics:
A:Gene: SGD:PHO81
A:Cross-references: SGD:S0003465; MIPS:YGR233C
A:Map position: 7R
C:Superfamily: ankryrin; ankryrin repeat homology
C:Keywords: transmembrane protein
F:330-346/Domain: transmembrane #status predicted <TM1>
F:904-920/Domain: transmembrane #status predicted <TM2>

Query Match 1.3%; Score 17; DB 2; Length 1178;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 742 NNNNNNNNNNNNNNNNNNNI 758

Db 233 NNNNNNNNNNNNNNNNNNNI 249

RESULT 10
T18287
protein-tyrosine kinase (EC 2.7.1.112) - slime mold (*Dictyostelium discoideum*)
C:Species: *Dictyostelium discoideum*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Feb-2000
C:Accession: T18287
R:Adler, K.
Submitted to the EMBL Data Library, July 1996
A:Reference number: Z18856
A:Accession: T18287
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1338 <ADL>
A:Cross-references: EMBL:U64830; NID:G1468982; PID:G1468983; PID:NAB04999.1
C:Genetics:
A:Introns: 1181/3
C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase

Query Match 1.3%; Score 17; DB 2; Length 1338;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 742 NNNNNNNNNNNNNNNNNNNI 758

Db 5 NNNNNNNNNNNNNNNNNNNI 21

RESULT 11
T14577
protein kinase YAKA (EC 2.7.1.-) - slime mold (*Dictyostelium discoideum*)
C:Species: *Dictyostelium discoideum*
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14577
R:Kuspa, A.; Lu, S.; Souza, G.M.
Submitted to the EMBL Data Library, January 1998
A:Description: YAKA, a protein kinase required for the growth to development transition
A:Reference number: Z18146
A:Accession: T14577
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1457 <KUS>
A:Cross-references: EMBL:AF045453; NID:G2854116; PID:G2854117; PID:AA02554.1
C:Genetics:
A:Gene: YAKA
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein ;

Query Match 1.3%; Score 17; DB 2; Length 1457;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 741 NNNNNNNNNNNNNNNNNNNI 757

Db 44 NNNNNNNNNNNNNNNNNNNI 60

RESULT 12
S71628
sensory transduction histidine kinase dokA - slime mold (*Dictyostelium discoideum*)
C:Species: *Dictyostelium discoideum*
C>Date: 27-Nov-1997 #sequence_revision 12-Dec-1997 #text_change 11-Jan-2002
C:Accession: S71628; S78068
R:Schuster, S.C.; Noegel, A.A.; Oehme, F.; Gerisch, G.; Simon, M.I.
EMBO J. 15, 3880-3889, 1996
A:Title: The hybrid histidine kinase DOKA is part of the osmotic response system of *Dictyostelium*
A:Reference number: S71628; MUID:96324396; PMID:8670893
A:Accession: S71628
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1670 <SCH>

A/Cross-references: EMBL:X96869
A/Experimental source: strain AX2; substrain 214
A/Schuster, S.C.; Noegel, A.A.; Oehme, F.; Gerisch, G.; Simon, M.I.
Submitted to the EMBL Data Library, March 1996
A/Description: The hybrid histidine kinase Doka is part of the osmotic response system C
A/Reference number: S78068
A/Molecule type: DNA
A/Accession: S78068
A/Residues: 1-149, 'E', 151-219, 'TRVLKLIQSTNNMIYWY', 238-1671 <SCW>
A/Cross-references: EMBL:X96869; NID:g1237201; PIDN:CAA65612.1; PID:g1237202
C/Genetics:
A/Gene: doka
C/Function:
A/Description: modulates cell response to changes in osmolarity; involved in spore forma
C/Superfamily: response regulator homology
C/Keywords: phosphoprotein; signal transduction
F:1520-1629/Domain: response regulator homology <RRH2>
F:1568/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 1.3%; Score 17; DB 2; Length 1671;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 741 TNNNNNNNNNNNNNNNN 757
DB 374 TNNNNNNNNNNNNNNNN 390

RESULT 13
T08875
histidine kinase homolog DHKB - slime mold (Dictyostelium discoideum)
N/Alternate names: hybrid histidine kinase DHKB
C/Species: Dictyostelium discoideum
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-May-2000
C/Accession: T08875
R:Zinda, M.J.; Singleton, C.K.
Dev. Biol. 196, 171-183, 1998
A/Title: The hybrid histidine kinase dhkb regulates spore germination in Dictyostelium d
A/Reference number: Z16506; MUID:98248997; PMID:9576830
A/Accession: T08875
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1969 <SIN>
A/Cross-references: EMBL:AF024654; NID:g2460282; PID:g2460283
A/Experimental source: strain KAX3
C/Genetics:
A/Gene: dhkb
A/Intons: 790/3
C/Superfamily: response regulator homology
C/Keywords: protein kinase; transmembrane protein
F:1841-1964/Domain: response regulator homology <RRH>

Query Match 1.3%; Score 17; DB 2; Length 1699;
Best Local Similarity 100.0%; Pred. No. 3.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNN 758
DB 37 NNNNNNNNNNNNNNNNN 53

RESULT 14
T18440
hypothetical protein C0425w - malaria parasite (Plasmodium falciparum)
C/Species: Plasmodium falciparum
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C/Accession: T18440
R:Lawson, D.; Bowman, S.; Barrell, B.
Submitted to the EMBL Data Library, August 1997
A/Reference number: Z18935
A/Accession: T18440
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA

A/Residues: 1-4550 <LAN>
A/Cross-references: EMBL:Z98547; NID:e1325376; PID:e1325396; PIDN:CAB11121.1
C/Genetics:
A/Map position: 3
A/Note: C0425w

Query Match 1.3%; Score 17; DB 2; Length 4550;
Best Local Similarity 100.0%; Pred. No. 6.8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNN 758
DB 1517 NNNNNNNNNNNNNNNNN 1533

RESULT 15
T02389
hypothetical protein At2g44340 (imported) - Arabidopsis thaliana
N/Alternate names: hypothetical protein F411.15
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C/Accession: T02389; C84877
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kau
Submitted to the EMBL Data Library, May 1998
A/Description: Arabidopsis thaliana chromosome II BAC F411 genomic sequence.
A/Reference number: Z14667
A/Accession: T02389
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-188 <ROU>
A/Cross-references: EMBL:AC004521; NID:g3128166; PID:g3128179
A/Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L
eus, D.; Nierman, W.C.; White, O.; Eisen, O.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: C84877
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-188 <STO>
A/Cross-references: GB:AE002093; NID:g3128179; PIDN:AAC16083.1; GSPDB:GM00139
C/Genetics:
A/Gene: F411.15; At2g44340
A/Map position: 2
C/Superfamily: Arabidopsis thaliana hypothetical protein T10K17.210

Query Match 1.2%; Score 16; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNN 757
DB 161 NNNNNNNNNNNNNNNNN 176

RESULT 16
S05356
hypothetical protein (clone pUK330) - slime mold (Dictyostelium discoideum) (fragment)
C/Species: Dictyostelium discoideum
C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 29-Oct-1999
C/Accession: S05356
R:Shaw, D.R.; Richter, H.; Glorda, R.; Ohmachi, T.; Emis, H.L.
Mol. Gen. Genet. 218, 453-459, 1989
A/Title: Nucleotide sequences of Dictyostelium discoideum developmentally regulated cdn.
A/Reference number: S05355; MUID:90066348; PMID:2511421
A/Accession: S05356
A/Molecule type: mRNA
A/Residues: 1-317 <SHA>
A/Cross-references: EMBL:X16523; NID:g7321; PIDN:CAA34530.1; PID:g7322

Query Match 1.2%; Score 16; DB 2; Length 317;

Best Local Similarity 100.0%; Pred. No. 6.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 742 NNNNNNNNNNNNNNNNN 757
|||||
Db 54 NNNNNNNNNNNNNNNNN 69

RESULT 17

T02046
prolamin box binding factor - maize
C/Species: Zea mays (maize)
C/Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 29-Oct-1999
C/Accession: T02046
R/Vicente-Carbajosa, J.; Moose, S.P.; Parsons, R.L.; Schmidt, R.J.
Proc. Natl. Acad. Sci. U.S.A. 94, 7685-7690, 1997
A/Title: A maize zinc-finger protein binds the prolamin box in zein gene promoters and
A/Reference number: 214511; MUID:97352860; PMID:9207153
A/Accession: T02046
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-328 <VIC>
A/Cross-references: EMBL:U82230; NID:G2393774; PIDN:AA870119.1; PID:G2393775
C/Genetics:
A/Gene: PBF

Query Match 1.2%; Score 16; DB 2; Length 328;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 742 NNNNNNNNNNNNNNNNN 757
|||||
Db 310 NNNNNNNNNNNNNNNNN 325

RESULT 18

T52593
squamosa promoter binding protein homolog 9 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 24-Oct-2000
C/Accession: T52593
R/Cardon, G.H.; Hohmann, S.; Nettesheim, K.; Saedler, H.; Huisser, P.
Plant J. 12, 367-377, 1997
A/Title: Functional analysis of the Arabidopsis thaliana SHP-box gene SPL3: a novel gene
A/Reference number: 224029; MUID:9301089; PMID:9301089
A/Accession: T52593
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-373 <CAR>
A/Cross-references: EMBL:AJ011639; PIDN:CAB56591.1
A/Experimental source: cultivar Landsberg erecta; Inflorescence
C/Genetics:
A/Gene: SPL3
A/Map position: 2
C/Keywords: DNA binding; transcription factor

Query Match 1.2%; Score 16; DB 2; Length 373;
Best Local Similarity 100.0%; Pred. No. 7.2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 742 NNNNNNNNNNNNNNNNN 757
|||||
Db 262 NNNNNNNNNNNNNNNNN 277

RESULT 19

T18285
hypochemical protein G2 - slime mold (Dictyostelium discoideum) plasmid
C/Species: Dictyostelium discoideum
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C/Accession: T18285; S28720
R/Reben Jr., W.K.; Gonzales, C.M.; Gonzales, S.T.; Pilkington, K.J.; Kiyosawa, H.; Hugh
Genetics 148, 1117-1125, 1998

A/Title: Dictyostelium discoideum nuclear plasmid Ddp5 is a chimera related to the Ddp1
A/Reference number: 214684; MUID:98198836; PMID:9539429

A/Accession: T18285
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-401 <RIE>

A/Cross-references: EMBL:U00796; NID:G2702254; PID:G932876; PIDN:AA18630.1
A/Experimental source: plasmid Ddp5
R/Gurniak, C.B.; Bang, A.G.; Noegel, A.A.
Curr. Genet. 17, 321-325, 1990

A/Title: Transcript and sequence analysis of a 5.1 kb contiguous fragment of Dictyostel
A/Reference number: S28718; MUID:90234839; PMID:2340592
A/Accession: S28720
A/Molecule type: DNA

A/Residues: 155-401 <GUR>
A/Cross-references: EMBL:X53237; NID:G7280; PIDN:CAA37329.1; PID:G7283

A/Experimental source: plasmid Ddp1
A/Note: the authors translated the codon TTT for residue 284 as Glu and GGT for residue
C/Genetics:
A/Gene: d-4
A/Genome: plasmid

Query Match 1.2%; Score 16; DB 2; Length 401;
Best Local Similarity 100.0%; Pred. No. 7.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 743 NNNNNNNNNNNNNNNNN 758
|||||
Db 358 NNNNNNNNNNNNNNNNN 373

RESULT 20
A46391
cAMP receptor subtype 3 - slime mold (Dictyostelium sp.)

C/Species: Dictyostelium sp.
C/Date: 18-Oct-1993 #sequence_revision 18-Nov-1994 #text_change 29-Oct-1999
C/Accession: A46391
R/Johnson, R.L.; Saxe III, C.L.; Gollop, R.; Kimmel, A.R.; Devreotes, P.N.
Genes Dev. 7, 273-282, 1993

A/Title: Identification and targeted gene disruption of cAR3, a cAMP receptor subtype e
A/Reference number: A46391; MUID:93170666; PMID:8382181
A/Accession: A46391
A/Status: preliminary
A/Molecule type: DNA

A/Residues: 1-490 <JOH>
A/Cross-references: GB:S55235; NID:G265733; PIDN:AA825437.1; PID:G265734
A/Note: sequence extracted from NCBI backbone (NCBIN:125367, NCBI:P:125369)

C/Keywords: cAMP binding; G protein-coupled receptor; phosphoprotein; transmembrane pro

Query Match 1.2%; Score 16; DB 2; Length 490;
Best Local Similarity 100.0%; Pred. No. 9.2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 742 NNNNNNNNNNNNNNNNN 757
|||||
Db 399 NNNNNNNNNNNNNNNNN 414

RESULT 21
S52830
HMS1 protein - yeast (Saccharomyces cerevisiae)

N/Alternate names: protein YM9916.09; protein YMR070w
C/Species: Saccharomyces cerevisiae
C/Date: 19-May-1995 #sequence_revision 01-Sep-1995 #text_change 29-Oct-1999

C/Accession: S52830; S59820
R/Pearson, D.; Bowman, S.
submitted to the EMBL Data Library, April 1995
A/Reference number: S52814
A/Accession: S52830
A/Molecule type: DNA

A/Residues: 1-490 <PEA>
A/Cross-references: EMBL:Z48952; NID:G763008; PIDN:CAA88795.1; PID:G763017; MIPS:YMR070

A/Experimental source: strain AB972

A:Reference number: 218935
A:Accession: T18446
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-600 <LAM>
A:Cross-references: EMBL:Z98547; NID:e1325376; PIDN:CAB51588.1
C:Genetics:
A:Map position: 3
A:Introns: 437/3; 477/2; 502/3; 532/3
A:Note: MAL3P3.11

Query Match 1.2%; Score 16; DB 2; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 742 NNNNNNNNNNNNNNNNN 757
Db 187 NNNNNNNNNNNNNNNNN 202

RESULT 26

A71607
Mtn3/RAC1P-like protein PFB0760w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession: A71607
R:Gardner, M.J.; Tetteelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: A71607
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-686 <GAR>
A:Cross-references: GB:AE001416; GB:AE001362; NID:g3845268; PIDN:AACT1941.1; PID:g384527
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0760w

Query Match 1.2%; Score 16; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 742 NNNNNNNNNNNNNNNNN 757
Db 50 NNNNNNNNNNNNNNNNN 65

RESULT 27

T51007
hypothetical protein B7J19.60 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T51007
R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T51007
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-720 <SCH>
A:Cross-references: EMBL:AL389899; GSPDB:GN00116; NCSP:B7J19.60
A:Experimental source: BAC clone B7J19; strain OR74A
C:Genetics:
A:Gene: NCSP:B7J19.60
A:Map position: 6
A:Introns: 534/2; 692/2

Query Match 1.2%; Score 16; DB 2; Length 720;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 742 NNNNNNNNNNNNNNNNN 757
Db 545 NNNNNNNNNNNNNNNNN 560

RESULT 28

S54522
hypothetical protein YMR164c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YMR520.13c
C:Species: Saccharomyces cerevisiae
C>Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 29-Oct-1999
C:Accession: S54522; S54609
R:Hunt, S.; Bowman, S.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54510
A:Accession: S54522
A:Molecule type: DNA
A:Residues: 1-758 <HUN>
A:Cross-references: GB:Z49705; EMBL:Z49700; NID:g825556; PIDN:CAA89800.1; PID:g825569;
A:Experimental source: strain AB972
C:Genetics:
A:Gene: SGD:MSS11
A:Cross-references: SGD:S0004774; MIPS:YMR164c
A:Map position: 13R

Query Match 1.2%; Score 16; DB 2; Length 758;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 742 NNNNNNNNNNNNNNNNN 757
Db 605 NNNNNNNNNNNNNNNNN 620

RESULT 29

T08611
hypothetical protein Doca - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C:Accession: T08611
R:Aubry, L.; Firtel, R.A.; Iranfar, N.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z16456
A:Accession: T08611
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-831 <AUB>
A:Cross-references: EMBL:AF020409; NID:g2425146; PID:g2425147
A:Experimental source: strain AX4
C:Genetics:
A:Gene: doca

Query Match 1.2%; Score 16; DB 2; Length 831;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 740 ITNNNNNNNNNNNNNNNN 755
Db 164 ITNNNNNNNNNNNNNNNN 179

RESULT 30

T08605
hypothetical protein He1f - slime mold (Dictyostelium discoideum) (fragment)
C:Species: Dictyostelium discoideum
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C:Accession: T08605
R:Loomis, W.F.; Iranfar, N.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z16451
A:Accession: T08605
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A;Residues: 1-947 <LOO>
A;Cross-references: EMBL:AF019881; NID:g2425112; PID:g2425113
A;Experimental source: strain AX4
C;Genetics:
A;Gene: helle

Query Match 1.2%; Score 16; DB 2; Length 947;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNN 757
DB 234 NNNNNNNNNNNNNNNN 249

RESULT 31

S67568
probable membrane protein YD035C - yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein D2749
C/Species: Saccharomyces cerevisiae

C/Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002
C/Accession: S67568

R;Paulin, L.; Saren, A.M.; Laamanen, P.
Submitted to the Protein Sequence Database, July 1996

A;Reference number: S67560

A;Accession: S67568
A;Molecule type: DNA

A;Residues: 1-961 <PAU>

A;Cross-references: EMBL:Z74083; NID:g1431014; PID:e252968; PID:g1431015; GSPDB:GN00004;
A;Experimental source: strain S288C

C;Genetics:

A;Gene: SGD:GPR1; MIPS:YD035C
A;Cross-references: SGD:S0002193

A;Map position: 4L

C;Keywords: transmembrane protein

F;72-73/Domain: transmembrane #status predicted <TM1>
F;72-108/Domain: transmembrane #status predicted <TM2>

F;140-156/Domain: transmembrane #status predicted <TM3>
F;181-197/Domain: transmembrane #status predicted <TM4>

F;256-272/Domain: transmembrane #status predicted <TM5>
F;623-639/Domain: transmembrane #status predicted <TM6>

F;662-678/Domain: transmembrane #status predicted <TM7>

Query Match 1.2%; Score 16; DB 2; Length 961;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNN 757

DB 515 NNNNNNNNNNNNNNNN 530

RESULT 32

JCS808
G protein-coupled receptor 1 - yeast (Saccharomyces cerevisiae)

C/Species: Saccharomyces cerevisiae
C/Date: 04-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 19-Apr-2002
C/Accession: JCS808

R;Fun, C.W.; Tanaki, H.; Nakayama, R.; Yamamoto, K.; Kumagai, H.
Biochem. Biophys. Res. Commun. 240, 287-292, 1997

A;Title: G-protein coupled receptor from yeast Saccharomyces cerevisiae.

A;Reference number: JCS808; MID:98049822; PMID:9388468
A;Accession: JCS808

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA
A;Residues: 1-962 <YUN>

C/Comment: This protein monitors the extracellular signal such as nutrition and transduc

C;Genetics:

A;Gene: SGD:GPR1; GPR1
A;Cross-references: SGD:S0002193

F;56-74/Domain: transmembrane #status predicted <TM1>
F;91-109/Domain: transmembrane #status predicted <TM2>

F;139-157/Domain: transmembrane #status predicted <TM3>

F;180-198/Domain: transmembrane #status predicted <TM4>
F;255-273/Domain: transmembrane #status predicted <TM5>
F;622-640/Domain: transmembrane #status predicted <TM6>
F;661-679/Domain: transmembrane #status predicted <TM7>

Query Match 1.2%; Score 16; DB 2; Length 962;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNN 757
DB 516 NNNNNNNNNNNNNNNN 531

RESULT 33

S48404
probable membrane protein YIL130W - yeast (Saccharomyces cerevisiae)

C/Species: Saccharomyces cerevisiae
C/Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 19-Apr-2002
C/Accession: S48404

R;Churcher, C.
Submitted to the EMBL Data Library, September 1994

A;Reference number: S48310
A;Accession: S48404
A;Molecule type: DNA

A;Residues: 1-964 <CHU>

A;Cross-references: GB:Z47047; EMBL:Z38059; NID:g603997; PID:g763216; MIPS:YIL130W
C;Genetics:

A;Cross-references: SGD:S0001392
A;Map position: 9L

C;Superfamily: unassigned GAL4-type zinc cluster proteins; GAL4 zinc binuclear cluster
C;Keywords: DNA binding; nucleas; transcription regulation; transmembrane protein

F;16-52/Domain: GAL4 zinc binuclear cluster homology <GLL4>
F;289-305/Domain: transmembrane #status predicted <TM1>

F;346-362/Domain: transmembrane #status predicted <TM2>

Query Match 1.2%; Score 16; DB 2; Length 964;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNN 757

DB 853 NNNNNNNNNNNNNNNN 868

RESULT 34

S54067
probable membrane protein YPR042C - yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein YP3085.06c; hypothetical protein YP9499.01c
C/Species: Saccharomyces cerevisiae
C/Date: 08-Jul-1995 #sequence_revision 26-Jul-1996 #text_change 19-Apr-2002
C/Accession: S54067; S61063

R;Padcock, K.; Churcher, C.M.
Submitted to the EMBL Data Library, May 1995

A;Reference number: S54059
A;Accession: S54067

A;Molecule type: DNA
A;Residues: 1-508 <BAD>

A;Cross-references: EMBL:Z49219; NID:g805025; PID:g805026; MIPS:YPR042C
A;Experimental source: strain AB972

R;Padcock, K.; Churcher, C.M.
Submitted to the EMBL Data Library, November 1995

A;Reference number: S61058
A;Accession: S61063

A;Molecule type: DNA
A;Residues: 465-1075 <BAW>

A;Cross-references: EMBL:Z68111; MIPS:YPR042C
C;Genetics:

A;Gene: SGD:PUP2
A;Cross-references: SGD:S0006246

A;Map position: 16R
C;Keywords: transmembrane protein

F;711-727/Domain: transmembrane #status predicted <TM1>

A;Gene: FlyBase:Su(z)2
A;Cross-references: FlyBase:FBgn0008654
A;Introns: 55/1; 139/2; 219/3
C;Superfamily: Drosophila suppressor protein of zeste; RING finger homology
C;Keywords: DNA binding; nucleus; zinc finger
F;31-79/Domain: RING finger homology <RING>

Query Match 1.2%; Score 16; DB 2; Length 1365;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNN 757
|||
Db 1077 NNNNNNNNNNNNNNNN 1092

RESULT 40

T14004
trfA protein - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C;Accession: T14004
R;Saigo, U.; Adachi, H.; Sutoh, K.
J. Biol. Chem. 273, 24654-24659, 1998
A;Title: Dictyostelium TRFA homologous to yeast Sen6 is required for normal growth and e
A;Reference number: Z17852; MID:98406112; PMID:9733762
A;Accession: T14004
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1390 <SAI>
A;Cross-references: EMBL:AB009080; NID:d1228566; PID:d1034109; PIDN:BAA33143.1
C;Genetics:
A;Gene: trfA
A;Introns: 333/3; 364/3; 637/1

Query Match 1.2%; Score 16; DB 2; Length 1390;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNN 757
|||
Db 830 NNNNNNNNNNNNNNNN 845

Search completed: October 2, 2003, 16:36:14
Job time : 78 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 2, 2003, 16:27:41 ; Search time 29 Seconds
(without alignments)
2088.634 Million cell updates/sec

Title: US-10-006-780-2

Perfect score: 1288
Sequence: 1 MNSKIKVVRKPLSELEKK.....KKLVQDNKSMDDNNHKK 1288

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	1.3	666	1	YEAF_YEAST
2	17	1.3	758	1	CC27_YEAST
3	17	1.3	1093	1	PI4K_DICDI
4	17	1.3	1178	1	PH81_YEAST
5	16	1.2	317	1	AAC4_DICDI
6	16	1.2	451	1	MYBH_DICDI
7	16	1.2	490	1	CAR3_DICDI
8	16	1.2	490	1	MOT3_YEAST
9	16	1.2	495	1	HSF1_ARATH
10	16	1.2	640	1	RAEP_CANAL
11	16	1.2	732	1	KMHB_DICDI
12	16	1.2	735	1	CIGB_DICDI
13	16	1.2	758	1	YMB8_YEAST
14	16	1.2	961	1	GPRI_YEAST
15	16	1.2	964	1	YINO_YEAST
16	16	1.2	989	1	PTP3_DICDI
17	16	1.2	1365	1	SUZ2_DROME
18	16	1.2	1584	1	KYK1_DICDI
19	16	1.2	1585	1	P3K3_DICDI
20	16	1.2	1743	1	TAGC_DICDI
21	16	1.2	1858	1	P3K2_DICDI
22	16	1.2	1905	1	TAGB_DICDI
23	15	1.2	451	1	ARBP_PLAFA
24	15	1.2	612	1	YP68_YEAST
25	15	1.2	837	1	CCRA_YEAST
26	15	1.2	888	1	YGB4_YEAST
27	15	1.2	2278	1	FABI_YEAST
28	14	1.1	700	1	CTK2_XENLA
29	14	1.1	730	1	NCD_DROME
30	14	1.1	730	1	KCM1_XENLA
31	14	1.1	745	1	ATK2_ARATH
32	14	1.1	754	1	ATK3_ARATH
33	14	1.1	793	1	ATK1_ARATH

34	14	1.1	832	1	KLPI_SCHPO
35	14	1.1	858	1	CYAG_DICDI
36	14	1.1	955	1	KINL_LEICH
37	14	1.1	1170	1	R115_YEAST
38	13	1.0	201	1	YMBD_YEAST
39	13	1.0	426	1	HKUB_LYCES
40	13	1.0	554	1	KLPI_SCHPO
41	13	1.0	669	1	FP51_YEAST
42	13	1.0	708	1	GBF1_DICDI
43	13	1.0	851	1	NUDI_YEAST
44	13	1.0	1151	1	GRI1_YEAST
45	12	0.9	212	1	YMF3_YEAST
46	12	0.9	537	1	ARP_PLAFA
47	12	0.9	590	1	HMF5_DROME
48	12	0.9	624	1	DSK1_CYLFU
49	12	0.9	647	1	KNRL_DROME
50	12	0.9	671	1	KP2C_MACFA
51	12	0.9	671	1	KP2C_RAT
52	12	0.9	679	1	KIF2_HUMAN
53	12	0.9	682	1	KIF2_XENLA
54	12	0.9	716	1	KIF2_MOUSE
55	12	0.9	718	1	KP2C_CRIGR
56	12	0.9	721	1	KP2C_MOUSE
57	12	0.9	725	1	KP2C_HUMAN
58	12	0.9	727	1	BRC1_DROME
59	12	0.9	914	1	AZP1_YEAST
60	12	0.9	1177	1	SP97_DICDI
61	12	0.9	1341	1	YL78_YEAST
62	12	0.9	1570	1	P3K1_DICDI
63	12	0.9	2452	1	RBP1_PLAFA
64	11	0.9	172	1	PKXA_DICDI
65	11	0.9	331	1	COL6_ARATH
66	11	0.9	443	1	CAR4_DICDI
67	11	0.9	448	1	AAC2_DICDI
68	11	0.9	472	1	HMCA_DROME
69	11	0.9	672	1	OSM3_CAEEL
70	11	0.9	673	1	KFC1_HUMAN
71	11	0.9	676	1	RLM1_YEAST
72	11	0.9	729	1	KAR3_YEAST
73	11	0.9	784	1	KLPI6_SCHPO
74	11	0.9	793	1	RECA_DICDI
75	11	0.9	805	1	KIP3_YEAST
76	11	0.9	935	1	KINH_SYNRA
77	11	0.9	987	1	ATK4_ARATH
78	11	0.9	1225	1	KF4A_CHICK
79	11	0.9	1226	1	KF4A_XENLA
80	11	0.9	1230	1	ST20_CANAL
81	11	0.9	1231	1	KF4A_MOUSE
82	11	0.9	1232	1	KF4A_HUMAN
83	11	0.9	2339	1	RPCI_PLAFA
84	10	0.8	240	1	MTNB_LISIN
85	10	0.8	240	1	MTNB_LISMO
86	10	0.8	368	1	YB11_YEAST
87	10	0.8	372	1	COUB_ARATH
88	10	0.8	522	1	IL9R_HUMAN
89	10	0.8	525	1	ESR1_MICUN
90	10	0.8	623	1	PMT1_DROME
91	10	0.8	648	1	KAPC_DICDI
92	10	0.8	718	1	PNT2_DROME
93	10	0.8	720	1	AB13_ARATH
94	10	0.8	726	1	NFI1_YEAST
95	10	0.8	823	1	SCH9_YEAST
96	10	0.8	865	1	E78A_DROME
97	10	0.8	972	1	OME_DROME
98	10	0.8	1038	1	KP17_MOUSE
99	10	0.8	1080	1	HDC_DROME
100	10	0.8	1195	1	YK76_YEAST
101	10	0.8	1228	1	ATC_PLAFA
102	10	0.8	1419	1	MDR_PLAFA
103	10	0.8	1584	1	U104_CAEEL
104	10	0.8	1648	1	KP14_HUMAN
105	10	0.8	1749	1	K13A_MOUSE
106	10	0.8	1805	1	K13A_HUMAN

Q92376	schizosach
Q03101	dictyostell
P46865	leishmania
P43565	saccharomyc
P38429	saccharomyc
Q22300	lycoperficc
Q94600	schizosach
P23300	saccharomyc
P36417	dictyostell
P32336	saccharomyc
P24814	saccharomyc
Q04978	saccharomyc
P04931	plasmodium
P07548	drosophila
Q39493	cylindrothe
P13054	drosophila
Q951P1	macaca fasc
Q62309	rattus norv
Q00139	homo sapien
Q91637	xenopus lae
P28740	mus musculu
P70096	cricketulus
Q92288	mus musculu
Q99661	homo sapien
Q01295	drosophila
P41696	saccharomyc
Q95293	dictyostell
Q05854	saccharomyc
P54673	dictyostell
P14248	plasmodium
Q23921	dictyostell
Q91427	arabidopsis
Q91493	dictyostell
P14196	dictyostell
P09085	drosophila
P46873	caenorhabd
Q96419	homo sapien
Q12224	saccharomyc
P17119	saccharomyc
Q59751	schizosach
Q23917	dictyostell
P53086	saccharomyc
Q43093	syncephalis
Q81335	arabidopsis
Q90640	gallus gall
Q91784	xenopus lae
Q92212	candida alb
P33174	homo sapien
Q95339	homo sapien
P27425	plasmodium
Q92459	listeria in
O84551	listeria mo
P38180	saccharomyc
Q95855	arabidopsis
Q01113	homo sapien
P57753	microsporont
P51022	drosophila
P34099	dictyostell
P51023	drosophila
Q01593	arabidopsis
Q12116	saccharomyc
P11992	saccharomyc
P45447	drosophila
Q24432	drosophila
Q996W8	mus musculu
Q924M8	drosophila
P36168	saccharomyc
Q08853	plasmodium
P13568	plasmodium
P23678	caenorhabd
Q15058	homo sapien
Q969W7	mus musculu
Q924H9	homo sapien

107	10	0.8	1826	1	K13B_HUMAN	Q9NQB8	homo sapien	180	8	0.6	397	1	YMP7_YEAST	Q04359	saccharomyc
108	9	0.7	375	1	CAR2_DICDI	P34907	dictyosteli	181	8	0.6	413	1	E2F4_HUMAN	O16254	homo sapien
109	9	0.7	377	1	PRP2_DICDI	P34138	dictyosteli	182	8	0.6	429	1	HML2_HELBO	O01622	helobdella
110	9	0.7	384	1	KFZ5_HUMAN	Q9U114	homo sapien	183	8	0.6	437	1	T1S1_DROME	P47980	drosophila
111	9	0.7	398	1	HKLI_ARATH	P46036	arabidopsis	184	8	0.6	443	1	YERO_YEAST	P39829	saccharomyc
112	9	0.7	403	1	GBA8_DICDI	P34063	dictyosteli	185	8	0.6	446	1	CHMO_BETUV	O22533	beta vulgar
113	9	0.7	416	1	PMO1_DICDI	P34100	dictyosteli	186	8	0.6	455	1	YB12_YEAST	P38288	saccharomyc
114	9	0.7	422	1	NTU9_YEAST	Q00199	saccharomyc	187	8	0.6	461	1	US45_LACLC	P22865	lactococcus
115	9	0.7	473	1	DCOR_CANAL	P78569	ratu	188	8	0.6	516	1	P54_ENTFC	P13622	enterococcu
116	9	0.7	508	1	EGRI_RAT	P08154	ratu	189	8	0.6	535	1	CAC3_DROME	Q9V427	drosophila
117	9	0.7	528	1	CTK1_YEAST	Q03957	saccharomyc	190	8	0.6	554	1	HAP4_YEAST	P14064	saccharomyc
118	9	0.7	564	1	EMRP_CANAL	P28873	candida alb	191	8	0.6	567	1	NUZM_HANWI	P48906	hansenula w
119	9	0.7	577	1	P2B1_DROME	P48456	drosophila	192	8	0.6	569	1	MANC_PIRSP	P55287	plumyces s
120	9	0.7	578	1	PSP2_YEAST	P50109	saccharomyc	193	8	0.6	571	1	NIR_PIRSP	P38500	betula verr
121	9	0.7	598	1	KLP3_CAEEL	P45962	caenorhabdi	194	8	0.6	583	1	NIR_BETVE	P21375	escherichia
122	9	0.7	658	1	UVRB_HELPJ	Q9KKA0	helicobacte	195	8	0.6	586	1	RT67_ECOLI	P50875	saccharomyc
123	9	0.7	658	1	UVRB_HELPY	P94846	helicobacte	196	8	0.6	604	1	SP20_YEAST	O15736	dictyosteli
124	9	0.7	675	1	UVRB_NEIGO	Q50939	neisseria g	197	8	0.6	612	1	T1PD_DICDI	P39511	saccharomyc
125	9	0.7	675	1	UVRB_NEIMA	P56996	neisseria m	198	8	0.6	653	1	DUS8_MOUSE	O09112	mus musculu
126	9	0.7	675	1	UVRB_NEIMB	O33395	neisseria m	199	8	0.6	669	1	UVRB_XYLFA	O9PEX1	xyloella fas
127	9	0.7	682	1	CONN_DROME	O01819	drosophila	200	8	0.6	669	1	UVRB_METAC	O8TK53	methanosarc
128	9	0.7	696	1	UVRB_RALSO	Q8Y0M2	raibetonia s	201	8	0.6	670	1	UVRB_METWA	O8PT29	methanosarc
129	9	0.7	713	1	YHC4_YEAST	P38741	saccharomyc	202	8	0.6	670	1	UVRB_XANAC	O8PJ51	xanthomonas
130	9	0.7	719	1	NRP1_YEAST	P32770	saccharomyc	203	8	0.6	673	1	UVRB_XANCP	O8PTX1	xanthomonas
131	9	0.7	740	1	KRE6_CANAL	P87023	candida alb	204	8	0.6	673	1	ECR_AEPAE	P49880	aedes aegypt
132	9	0.7	749	1	MAD1_YEAST	P40957	saccharomyc	205	8	0.6	675	1	SIM_DROME	P05709	drosophila
133	9	0.7	770	1	KLP4_EMENT	P28739	emeritella	206	8	0.6	697	1	ARA_DROME	O24248	drosophila
134	9	0.7	779	1	SRP_DROME	P52172	drosophila	207	8	0.6	717	1	KL68_DROME	P46687	drosophila
135	9	0.7	782	1	SP4_MOUSE	Q62445	mus musculu	208	8	0.6	784	1	STUB_DROME	O05319	drosophila
136	9	0.7	784	1	SP4_HUMAN	O02446	homo sapien	209	8	0.6	786	1	MP75_YEAST	P39016	saccharomyc
137	9	0.7	868	1	NRG2_RAT	O35569	ratu	210	8	0.6	834	1	UBPA_DICDI	P54201	dictyosteli
138	9	0.7	904	1	ABRU_DROME	O24174	drosophila	211	8	0.6	837	1	PSPI_YEAST	P50896	saccharomyc
139	9	0.7	928	1	PRP3_YEAST	P40048	saccharomyc	212	8	0.6	841	1	R132_ARATH	O9SIE5	arabidopsis
140	9	0.7	960	1	DIG1_DROME	P31007	drosophila	213	8	0.6	847	1	LAC9_KUTLA	P08657	kluyveromyc
141	9	0.7	1000	1	S155_YEAST	P43612	saccharomyc	214	8	0.6	865	1	CLPC_CHLMU	O9PKC8	chlamydia m
142	9	0.7	1091	1	YB01_YEAST	P47135	saccharomyc	215	8	0.6	870	1	TRK2_SCHPO	Q10065	schizosacch
143	9	0.7	1094	1	YB00_YEAST	P38114	saccharomyc	216	8	0.6	880	1	KLP5_SCHPO	O14343	schizosacch
144	9	0.7	1188	1	OSH1_YEAST	P35845	saccharomyc	217	8	0.6	883	1	G2D1_HUMAN	Q9U119	h. general t
145	9	0.7	1227	1	JIP3_DROME	O9GQF5	drosophila	218	8	0.6	959	1	GUNA_PSEFL	P10476	pseudomonas
146	9	0.7	1398	1	TOP2_PLAFK	P41001	plasmodium	219	8	0.6	962	1	DA81_YEAST	P21657	saccharomyc
147	9	0.7	1407	1	CYAA_DICDI	Q03160	dictyosteli	220	8	0.6	970	1	Y277_MYCGB	Q49409	mycoplasma
148	9	0.7	1556	1	PRO5_DROI	O9A6A1	drosophila	221	8	0.6	970	1	CRM_DROSE	O8MX88	drosophila
149	9	0.7	1677	1	VIT_ACTR	O90243	acapsenser t	222	8	0.6	975	1	MGR_DROME	P91655	drosophila
150	9	0.7	3092	1	IRAI_YEAST	P18963	saccharomyc	223	8	0.6	976	1	G2D1_MOUSE	P31521	drosophila
151	8	0.6	115	1	SER1_GALME	O96614	galleria me	224	8	0.6	1028	1	FTF1_DROME	P31244	drosophila
152	8	0.6	117	1	DHA_CARPL	P22238	catelostig	225	8	0.6	1043	1	SGG_DROME	P18411	drosophila
153	8	0.6	121	1	RS6_RICCN	Q92JK3	ricicetstia	226	8	0.6	1067	1	Y124_METUA	Q73568	methanococc
154	8	0.6	121	1	RS6_RICCN	Q92EAE	ricicetstia	227	8	0.6	1075	1	PLC1_CANAL	O13453	candida alb
155	8	0.6	150	1	ASNC_HAEIN	P44337	haemophilus	228	8	0.6	1099	1	G2D1_MOUSE	Q9J157	mus musculu
156	8	0.6	193	1	UBC6_HUMAN	P51965	homo sapien	229	8	0.6	1104	1	IRSI_HUMAN	P35568	homo sapien
157	8	0.6	193	1	UBC6_MOUSE	P52482	mus musculu	230	8	0.6	1242	1	SMI1_YEAST	P09547	saccharomyc
158	8	0.6	215	1	PRF2_RAT	P70492	ratu	231	8	0.6	1314	1	IRSI_MOUSE	P81122	mus musculu
159	8	0.6	220	1	SER2_GALME	O96615	galleria me	232	8	0.6	1321	1	IRSI_MOUSE	Q9Y412	homo sapien
160	8	0.6	260	1	DH10_ARATH	P42759	arabidopsis	233	8	0.6	1324	1	PRO5_DROME	P29671	drosophila
161	8	0.6	265	1	JOIN_LYCPS	P31168	arabidopsis	234	8	0.6	1403	1	N153_RAT	P49791	ratu
162	8	0.6	265	1	JOIN_LYCPS	O9FYU6	lycopersico	235	8	0.6	1468	1	N153_RAT	P49790	homo sapien
163	8	0.6	284	1	SMX5_SCHMA	Q26604	schistosoma	236	8	0.6	1475	1	N153_RAT	O04182	saccharomyc
164	8	0.6	288	1	KADC_SOLTU	Q8H5W1	solanum tub	237	8	0.6	1529	1	PDRF_YEAST	O04182	saccharomyc
165	8	0.6	298	1	MYOD_CHICK	P16075	gallus gall	238	8	0.6	1597	1	SOL_DROME	P27398	drosophila
166	8	0.6	329	1	CAHX_FIAPR	P46281	flaveria pr	239	8	0.6	1679	1	FUR2_DROME	P30452	drosophila
167	8	0.6	330	1	CAH1_FIALI	P46512	flaveria pr	240	8	0.6	1704	1	VIT1_FUNHE	Q90508	funulus he
168	8	0.6	330	1	CAHX_FIAPR	P46511	flaveria pr	241	8	0.6	1850	1	VIT2_CHICK	P02848	gallus gall
169	8	0.6	331	1	CAHX_FIAPR	P46511	flaveria pr	242	8	0.6	2175	1	HMCU_DROME	P01810	drosophila
170	8	0.6	334	1	YPO6_CAEEL	P55112	caenorhabdi	243	8	0.6	3005	1	ZFH2_DROME	P28167	drosophila
171	8	0.6	339	1	RMAR_CANGA	P21358	candida gla	244	8	0.6	3726	1	TRX_DROME	P20659	drosophila
172	8	0.6	354	1	URE2_YEAST	P23202	saccharomyc	245	8	0.5	46	1	VIT_DROME	P56530	anas platyr
173	8	0.6	361	1	CGD2_ARATH	P42752	arabidopsis	246	7	0.5	70	1	VNCH_ECOLI	P61118	escherichia
174	8	0.6	367	1	CSP_PLAYO	P06914	plasmodium	247	7	0.5	91	1	MINE_WIGER	Q66211	wiggleswort
175	8	0.6	369	1	DAG2_ARATH	O9P9Y0	arabidopsis	248	7	0.5	92	1	HMG2_TERTH	P40627	tetrahymena
176	8	0.6	372	1	MSSP_HUMAN	O60909	h. beta-1, 4-	249	7	0.5	97	1	VN13_MERTU	P25769	methanococc
177	8	0.6	372	1	MSSP_HUMAN	P23558	homo sapien	250	7	0.5	98	1	HMEN_HELTR	P23397	helobdella
178	8	0.6	396	1	RMAR_YEAST	P02381	saccharomyc	251	7	0.5	99	1	Y010_BPHPI	P51712	bacteriophen
179	8	0.6	397	1	ODPA_PEA	P52902	pleum sativ	252	7	0.5	100	1	LGI_TETTH	P11873	tetrahymena

253	7	0.5	102	1	109_LYCES	Q43495 lycopersico	326	7	0.5	247	1	DB83_HUMAN	P57088 homo sapien
254	7	0.5	102	1	M1H_PENVA	P55322 penaeus van	327	7	0.5	247	1	DB83_MOUSE	Q9c167 mus musculu
255	7	0.5	104	1	YJ04_YEAST	P47009 saccharomyc	328	7	0.5	247	1	DB83_RAT	Q92142 rattus norv
256	7	0.5	112	1	VE4_HPV29	P51897 human papil	329	7	0.5	250	1	UBC6_YEAST	P13396 saccharomyc
257	7	0.5	115	1	CTF9_MOUSE	Q9c177 mus musculu	330	7	0.5	251	1	YCS2_METUA	Q58448 methanococ
258	7	0.5	117	1	YFB9_YEAST	P43576 saccharomyc	331	7	0.5	251	1	LI39_CAEEL	P34684 caenorhabdi
259	7	0.5	122	1	M126_CHICK	P28318 gallus gall	332	7	0.5	254	1	EMD_HUMAN	P50402 homo sapien
260	7	0.5	129	1	PA25_TRIST	P82896 trimeureuru	333	7	0.5	254	1	S5A2_MOUSE	Q99169 mus musculu
261	7	0.5	123	1	RL7_WIGBR	Q8d234 wigglewort	334	7	0.5	255	1	BDNF_CAVPO	O70183 cavia porce
262	7	0.5	124	1	YBF3_YEAST	P38190 saccharomyc	335	7	0.5	255	1	HXD4_HUMAN	P09016 homo sapien
263	7	0.5	128	1	DHX1_ARATH	P25863 arabidopsis	336	7	0.5	258	1	VC23_VACCL	P19663 vaccinia vl
264	7	0.5	130	1	Y816_ARATH	Q81868 arabidopsis	337	7	0.5	259	1	YG31_YEAST	P19369 saccharomyc
265	7	0.5	139	1	DH1_HORVU	P12951 hordeum vul	338	7	0.5	260	1	OLG1_MOUSE	Q91xm5 mus musculu
266	7	0.5	143	1	DH2_HORVU	P12952 hordeum vul	339	7	0.5	260	1	ZFP4_ARATH	Q39263 arabidopsis
267	7	0.5	145	1	DH11_GOSHI	P09442 gosyrylum h	340	7	0.5	262	1	COR4_WHEAT	P46524 triticum ae
268	7	0.5	148	1	PF02_ARATH	Q91398 arabidopsis	341	7	0.5	263	1	COL1_PANCA	P11885 rana catesb
269	7	0.5	149	1	DH15_WHEAT	Q00742 triticum ae	342	7	0.5	263	1	PSF1_ARATH	Q42029 arabidopsis
270	7	0.5	150	1	ZFP2_ARATH	Q39261 arabidopsis	343	7	0.5	265	1	PSF2_ARATH	Q49344 arabidopsis
271	7	0.5	151	1	DH1D_ORYSA	P22293 oryza sativ	344	7	0.5	265	1	Y1F6_YEAST	P40572 saccharomyc
272	7	0.5	155	1	DHB_CRAVL	P22239 cratereotig	345	7	0.5	266	1	SUE1_PYRAE	Q8zu79 pyrobaculum
273	7	0.5	160	1	KF1C_MOUSE	Q35071 mus musculu	346	7	0.5	268	1	LB13_ARATH	Q9a61 arabidopsis
274	7	0.5	160	1	NUSB_CHLPN	Q92660 chlamydia p	347	7	0.5	268	1	UPK_BUCBP	P59523 buchnera ap
275	7	0.5	161	1	DH3_HORVU	P12948 hordeum vul	348	7	0.5	273	1	FMDG_METUA	Q58571 methanococ
276	7	0.5	163	1	DH21_ORYSA	P12253 oryza sativ	349	7	0.5	274	1	PSBS_SEITOL	Q02060 splinacea ol
277	7	0.5	164	1	DH1B_ORYSA	P22911 oryza sativ	350	7	0.5	275	1	HXA5_HETFR	Q91a23 heterodontu
278	7	0.5	164	1	DH1C_ORYSA	P22912 oryza sativ	351	7	0.5	280	1	V1E3_HCMVT	P06434 human cytom
279	7	0.5	166	1	YEF4_YEAST	P32617 saccharomyc	352	7	0.5	284	1	CAS1_RAT	P02661 rattus norv
280	7	0.5	168	1	DH1_MAIZE	P12950 zea mays (m	353	7	0.5	285	1	Y007_BORBU	O51040 borrelia bu
281	7	0.5	168	1	PIG1_DROME	P26023 drosophila	354	7	0.5	285	1	YK62_CAEEL	P46552 caenorhabdi
282	7	0.5	169	1	Y358_BUCAL	P57439 buchnera ap	355	7	0.5	292	1	Y042_CAEEL	P34680 caenorhabdi
283	7	0.5	170	1	COP1_ARATH	Q39065 arabidopsis	356	7	0.5	297	1	TX20_MOUSE	Q9e603 mus musculu
284	7	0.5	170	1	DEF_AGR75	Q8u1d1 agrobacteri	357	7	0.5	304	1	WR13_ARATH	Q9srb7 arabidopsis
285	7	0.5	170	1	HSB7_HUMAN	Q9ubv9 homo sapien	358	7	0.5	306	1	EC40_DAUCA	Q07322 danus caro
286	7	0.5	174	1	DEF_RHIME	Q92sh6 rhizobium m	359	7	0.5	308	1	ZDH7_HUMAN	Q9uxf8 homo sapien
287	7	0.5	184	1	DHLE_RAPSA	P21298 raphanus sa	360	7	0.5	310	1	UDP_HUMAN	Q16631 homo sapien
288	7	0.5	185	1	DH14_ARATH	P42763 arabidopsis	361	7	0.5	313	1	TF62_DROME	P20332 drosophila
289	7	0.5	188	1	ERBP_RAT	P06911 rattus norv	362	7	0.5	313	1	YMS8_YEAST	Q03595 saccharomyc
290	7	0.5	190	1	ARF_GIALA	P26991 giardia lam	363	7	0.5	314	1	SINA_DROME	P23601 drosophila
291	7	0.5	193	1	LB12_ARATH	Q81bw3 arabidopsis	364	7	0.5	314	1	SINA_DROVI	P29904 drosophila
292	7	0.5	194	1	YK42_PYRPU	Q8tze3 pyrococcus	365	7	0.5	323	1	OTX1_BRARE	O91194 brachydantio
293	7	0.5	198	1	AMOS_DROME	Q9y087 drosophila	366	7	0.5	324	1	YL10_ARCFU	Q28112 archaeoglob
294	7	0.5	198	1	C4_GIALA	Q01832 giardia lam	367	7	0.5	325	1	YH03_YEAST	P38844 saccharomyc
295	7	0.5	198	1	CASI_CAVPO	P40656 cavia porce	368	7	0.5	328	1	CAHC_PEA	P17067 pisum sativ
296	7	0.5	200	1	HMGH_STRPU	P40644 strongyloce	369	7	0.5	328	1	DLX2_HUMAN	Q07687 homo sapien
297	7	0.5	203	1	YGBR_YEAST	P53288 saccharomyc	370	7	0.5	329	1	PS7B_MYCGE	P47850 mycoplasma
298	7	0.5	207	1	NUGM_HANWI	P48927 hansenula w	371	7	0.5	333	1	PIX2_CHICK	O93385 gallus gall
299	7	0.5	207	1	YUT2_YEAST	P47087 saccharomyc	372	7	0.5	334	1	LYTE_BACSU	P54421 bacillus su
300	7	0.5	211	1	SOCI_HUMAN	O15524 homo sapien	373	7	0.5	335	1	HXDD_HUMAN	P35453 homo sapien
301	7	0.5	212	1	SOCI_MOUSE	Q35716 mus musculu	374	7	0.5	336	1	CAHC_ARATH	P27140 arabidopsis
302	7	0.5	213	1	SOCI_RAT	Q9qx78 rattus norv	375	7	0.5	338	1	FEN_METAC	O8tiy5 methanosarc
303	7	0.5	213	1	HPI_DROVI	P29227 drosophila	376	7	0.5	338	1	FEN_METMA	O8y1f6 methanosarc
304	7	0.5	216	1	ANSE_EIMTE	P15744 eimeria ten	377	7	0.5	338	1	HRCA_THENA	Q9wzvs thermotoga
305	7	0.5	218	1	ESME_DROME	O97177 drosophila	378	7	0.5	339	1	OTX3_BRARE	Q90677 brachydantio
306	7	0.5	218	1	WR12_ARATH	Q93wcy4 arabidopsis	379	7	0.5	339	1	HXDD_MOUSE	P70217 mus musculu
307	7	0.5	219	1	COL4_TTV1	P19273 thermoprote	380	7	0.5	344	1	TRPD_BRUME	O8y1f7 bruceella me
308	7	0.5	222	1	COL1_THUOB	Q9y0k2 t corticocor	381	7	0.5	349	1	YB1E_SCHPO	P81817 catharanthu
309	7	0.5	223	1	DEOC_STRP3	O8nu69 streptococc	382	7	0.5	347	1	E13A_SOYBN	O03173 glycine max
310	7	0.5	223	1	DEOC_STRPY	O99y51 streptococc	383	7	0.5	347	1	UTR2_YEAST	P32623 saccharomyc
311	7	0.5	224	1	RM01_RECAM	O21235 reclinomona	384	7	0.5	347	1	V1T3_CHICK	Q91025 gallus gall
312	7	0.5	225	1	DH4_HORVU	P12949 hordeum vul	385	7	0.5	352	1	STSY_CATRO	P18417 catharanthu
313	7	0.5	225	1	DH25_ORYSA	P30287 oryza sativ	386	7	0.5	355	1	HKL6_LYCES	O22299 lycopersico
314	7	0.5	229	1	COAT_TAV	P23627 tomato aspe	387	7	0.5	355	1	YDP6_SCHPO	O14009 schizosach
315	7	0.5	229	1	LB11_ARATH	Q96k08 arabidopsis	388	7	0.5	357	1	V136_DICDI	Q94480 dictyostell
316	7	0.5	230	1	DAG_ANTMA	Q38732 antirrhinum	389	7	0.5	363	1	CYX2_RHIME	O52915 rhizobium m
317	7	0.5	235	1	YH84_YEAST	P38727 saccharomyc	390	7	0.5	364	1	NK61_MESAU	Q60554 mesocricetu
318	7	0.5	238	1	CW14_YEAST	O13147 saccharomyc	391	7	0.5	365	1	NK61_MOUSE	Q99ma9 mus musculu
319	7	0.5	242	1	MNBA_MAIZE	P38564 zea mays (m	392	7	0.5	365	1	GGPE_SINNL	O43133 sinapis alb
320	7	0.5	242	1	NAGB_MYCPE	O8ewm7 mycoplasma	393	7	0.5	367	1	Y797_METUA	P58407 methanococ
321	7	0.5	243	1	ERP2_ARATH	O80338 arabidopsis	394	7	0.5	367	1	NK61_HUMAN	O58207 methanococ
322	7	0.5	243	1	IM17_ARATH	O9sp35 arabidopsis	395	7	0.5	368	1	ROX1_YEAST	P25042 saccharomyc
323	7	0.5	244	1	VC23_VACCC	P21090 vaccinia vi	396	7	0.5	369	1	MAGA_HUMAN	P43363 homo sapien
324	7	0.5	244	1	YENR_YEREN	P54295 yerlinala en	397	7	0.5	372	1	COO2_YEAST	P32378 saccharomyc
325	7	0.5	246	1	AX1B_ARATH	Q38829 arabidopsis	398	7	0.5	372	1		

399	7	0.5	375	1	OTC_PEA	043814	pisum sativ	472	7	0.5	472	1	K1CP_HUMAN	P08779	homo sapien
400	7	0.5	376	1	YBC3_YEAST	P38201	saccharomyc	473	7	0.5	473	1	PLSB_SPIOL	Q43689	spiniacia ol
401	7	0.5	377	1	VE2_HPV13	Q02263	human papil	474	7	0.5	474	1	ATH1_ARATH	P48731	arabidopsis
402	7	0.5	377	1	VMAT_P12HT	P24266	human para1	475	7	0.5	474	1	LAM3_MOUSE	P48660	mus musculu
403	7	0.5	377	1	Y835_METUA	Q58245	mechanococ	476	7	0.5	474	1	SOX4_HUMAN	Q06945	homo sapien
404	7	0.5	378	1	KLP2_BOMMO	P48245	bombyx mori	477	7	0.5	475	1	SIM1_YEAST	P40472	saccharomyc
405	7	0.5	379	1	HMB1_SOYBN	P46608	glycine max	478	7	0.5	476	1	CBP2_HORVU	P08818	hordium vul
406	7	0.5	382	1	TP6A_PIRFU	Q9v131	pyrococcus	479	7	0.5	476	1	REF3_SACBA	P05512	saccharomyc
407	7	0.5	382	1	TP6A_PIRFU	Q80K03	pyrococcus	480	7	0.5	479	1	DBP1_BACSU	P42305	bacillus su
408	7	0.5	382	1	TP6A_PIRFU	Q59209	pyrococcus	481	7	0.5	479	1	YP66_YEAST	Q12194	saccharomyc
409	7	0.5	383	1	GIC2_YEAST	Q06648	saccharomyc	482	7	0.5	481	1	HH_DROH1	P56674	drosophila
410	7	0.5	383	1	TRB2_ARATH	Q39242	arabidopsis	483	7	0.5	482	1	YSDR_CAEEL	Q09950	caenorhabdi
411	7	0.5	386	1	RMAR_HAWNI	P48849	hansenuia w	484	7	0.5	484	1	YORO_YEAST	Q08193	saccharomyc
412	7	0.5	387	1	WR36_ARATH	Q9at47	arabidopsis	485	7	0.5	485	1	Y136_TREPA	Q83112	treponema p
413	7	0.5	389	1	SERI_BOMMO	P07856	bombyx mori	486	7	0.5	487	1	OAF_DROME	Q9n166	drosophila
414	7	0.5	392	1	ODD_DROME	P23803	drosophila	487	7	0.5	489	1	MDM2_MOUSE	P23804	mus musculu
415	7	0.5	393	1	IRTF_HUMAN	Q00978	homo sapien	488	7	0.5	489	1	SFR4_MOUSE	Q8ve17	mus musculu
416	7	0.5	394	1	VATC_SCHPO	Q9hdw6	schizosacch	489	7	0.5	489	1	WR47_ARATH	Q9ze17	arabidopsis
417	7	0.5	397	1	HKL2_MALDO	Q04135	malus domes	490	7	0.5	491	1	KCS3_HUMAN	Q9bq31	homo sapien
418	7	0.5	397	1	YK08_CAEEL	P34303	caenorhabdi	491	7	0.5	491	1	KCS3_RABIT	Q9ct17	oryctolagus
419	7	0.5	398	1	CAR4_RHINI	Q03700	rhizopus ni	492	7	0.5	491	1	KCS3_RAT	Q88759	rattus norv
420	7	0.5	399	1	HM39_CAEEL	Q22812	caenorhabdi	493	7	0.5	493	1	HMA8_DROME	P09087	drosophila
421	7	0.5	400	1	OMPA_THEMEA	P25724	thermotoga	494	7	0.5	500	1	DCE_PETHY	Q07346	petunia hyb
422	7	0.5	401	1	NNOS_DROME	P32583	saccharomyc	495	7	0.5	501	1	ANKH_BRARE	P58368	brachydanio
423	7	0.5	406	1	SR40_YEAST	Q16186	homo sapien	496	7	0.5	503	1	WSC2_YEAST	P26536	saccharomyc
424	7	0.5	407	1	ADRM_HUMAN	Q16186	homo sapien	497	7	0.5	504	1	VLI1_HPV51	P26536	human papil
425	7	0.5	407	1	ADRM_MOUSE	Q91kvl	mus musculu	498	7	0.5	505	1	GPMT_MYCPU	P98947	mycoplasma
426	7	0.5	407	1	ADRM_MOUSE	Q91kvl	mus musculu	499	7	0.5	506	1	NPL3_HUMAN	Q99447	homo sapien
427	7	0.5	407	1	PE11_YEAST	P23250	saccharomyc	500	7	0.5	508	1	CROC_DROME	P32027	drosophila
428	7	0.5	410	1	PGK_METFE	P20971	methanother	501	7	0.5	510	1	KPK2_PLAFK	Q02555	plasmodium
429	7	0.5	410	1	PO42_HUMAN	Q12837	homo sapien	502	7	0.5	510	1	MS2P_CRIGR	Q54862	cricetulus
430	7	0.5	410	1	VIE2_HCMWT	P06437	human cytom	503	7	0.5	511	1	GUNB_PSEFL	P18116	pseudomonas
431	7	0.5	411	1	PO42_MOUSE	Q63934	mus musculu	504	7	0.5	512	1	GLGT_VICFA	Q04921	vicia faba
432	7	0.5	411	1	VIE2_HCMVA	P19893	human cytom	505	7	0.5	512	1	WR33_ARATH	Q88595	arabidopsis
433	7	0.5	411	1	VA12_SCHPO	Q09685	schizosacch	506	7	0.5	512	1	COX1_PAPHA	Q92x72	papio hamad
434	7	0.5	414	1	NSR1_YEAST	P27476	saccharomyc	507	7	0.5	513	1	CBX2_MOUSE	P30658	mus musculu
435	7	0.5	416	1	SO_DROME	Q9v3q2	drosophila	508	7	0.5	519	1	PTP1_DICDI	Q43422	homo sapien
436	7	0.5	421	1	SYT1_MOUSE	Q27350	drosophila	509	7	0.5	519	1	PTP1_DICDI	P34137	dictyosteli
437	7	0.5	421	1	SYT1_MOUSE	P46096	mus musculu	510	7	0.5	521	1	HME2_SCHMA	Q26601	schistosoma
438	7	0.5	421	1	SYT1_MOUSE	P21707	rattus norv	511	7	0.5	524	1	HUNB_TRICA	Q01791	tribolium c
439	7	0.5	422	1	SYT1_BOVIN	P46018	bos taurus	512	7	0.5	524	1	YACB_YEAST	P38712	saccharomyc
440	7	0.5	422	1	SYT1_HUMAN	P21579	homo sapien	513	7	0.5	528	1	ICP0_HAVER	P28990	equine heip
441	7	0.5	423	1	VDR_RAT	P13053	rattus norv	514	7	0.5	532	1	YEN1_SCHPO	P13695	echinosacch
442	7	0.5	424	1	MYCN_CHICK	P47191	gallus gall	515	7	0.5	536	1	SP70_DICDI	P52629	dictyosteli
443	7	0.5	427	1	MYCN_SERCA	P24504	serinus can	516	7	0.5	537	1	IFC3_EUGER	P51677	euglena gra
444	7	0.5	427	1	ST61_DISOM	P24504	serinus can	517	7	0.5	538	1	EGRI_HUMAN	P18116	homo sapien
445	7	0.5	429	1	DR48_YEAST	P18899	saccharomyc	518	7	0.5	543	1	HMA1_ARATH	P42804	arabidopsis
446	7	0.5	431	1	PO23_MOUSE	P31362	mus musculu	519	7	0.5	543	1	RRP3_YEAST	P38712	saccharomyc
447	7	0.5	433	1	AFIR_EMENI	P52957	emeritella	520	7	0.5	543	1	THSB_THEAC	P48425	thermoplasm
448	7	0.5	433	1	RTCL_DICDI	Q15746	dictyosteli	521	7	0.5	543	1	PHR2_CANAL	Q13318	candida alb
449	7	0.5	434	1	VO23_ROMPV	Q915h8	fowlpox vir	522	7	0.5	546	1	ERF_HUMAN	Q91x58	homo sapien
450	7	0.5	440	1	SOX4_MOUSE	Q06831	mus musculu	523	7	0.5	548	1	WR72_ARATH	Q91x58	arabidopsis
451	7	0.5	440	1	VRN2_ARATH	Q8w5b1	arabidopsis	524	7	0.5	548	1	EPD2_CANMA	Q74117	candida mal
452	7	0.5	440	1	FXF2_HUMAN	Q12947	homo sapien	525	7	0.5	549	1	ERF_MOUSE	P70459	mus musculu
453	7	0.5	449	1	TIG_TREPA	Q9pge4	ureaplasma	526	7	0.5	551	1	YGIF_YEAST	P53214	saccharomyc
454	7	0.5	450	1	THI2_YEAST	P38141	saccharomyc	527	7	0.5	552	1	HMEN_DROME	P02836	drosophila
455	7	0.5	450	1	UTH1_YEAST	P36135	saccharomyc	528	7	0.5	552	1	ENS2_YEAST	P12294	saccharomyc
456	7	0.5	452	1	K1CO_MOUSE	Q64144	mus musculu	529	7	0.5	559	1	GAS1_YEAST	P22146	saccharomyc
457	7	0.5	452	1	PUB1_YEAST	P31588	saccharomyc	530	7	0.5	559	1	GAT2_YEAST	Q40209	saccharomyc
458	7	0.5	453	1	MDHP_FLABI	P46489	flaveria bi	531	7	0.5	560	1	P2B2_DROME	Q40954	candida alb
459	7	0.5	461	1	PLSB_PHAVU	Q43822	phaseolus v	532	7	0.5	567	1	CH13_CANAL	P42568	homo sapien
460	7	0.5	462	1	SELA_CLOPE	Q8x1k2	clostridium	533	7	0.5	567	1	AP9_HUMAN	P57717	brachydanio
461	7	0.5	464	1	CAP_DICDI	P54654	dictyosteli	534	7	0.5	569	1	ESR1_BRARE	P23031	pseudomonas
462	7	0.5	464	1	IFB_BRALA	Q04948	branchiosteo	535	7	0.5	571	1	XYNC_PSEFL	P40095	saccharomyc
463	7	0.5	465	1	HNF6_HUMAN	Q9ubdc0	homo sapien	536	7	0.5	573	1	Y8Y8_YEAST	Q61909	mus musculu
464	7	0.5	465	1	HNF6_MOUSE	Q08755	mus musculu	537	7	0.5	577	1	MTG8_MOUSE	Q89b09	buchnera ap
465	7	0.5	466	1	CYP8_RAT	P50116	caenorhabdi	538	7	0.5	578	1	M3X7_MOUSE	Q62073	mus musculu
466	7	0.5	466	1	KG3H_DICDI	P51136	dictyosteli	539	7	0.5	580	1	MDLB_BUCAP	Q8k984	buchnera ap
467	7	0.5	468	1	Y134_CAEEL	P34425	caenorhabdi	540	7	0.5	580	1	SYN2_HUMAN	Q92777	homo sapien
468	7	0.5	469	1	EC33_YEAST	P38248	saccharomyc	541	7	0.5	582	1	AXU1_MOUSE	P59054	mus musculu
469	7	0.5	470	1	PLSB_CUCSA	Q39639	cucumis sat	543	7	0.5	583	1	CH12_CANAL	P40953	candida alb
470	7	0.5	471	1	HH_DROME	Q02936	drosophila	544	7	0.5	583	1			

545	7	0.5	586	1	HM26_CABEL	P34522	caenorhabdi	618	7	0.5	689	1	KFB1_RAT	O88658	rattus norv
546	7	0.5	586	1	SYN2_RAT	O63537	rattus norv	619	7	0.5	690	1	GSHP_CANAL	O94718	candida alb
547	7	0.5	589	1	AXU1_HUMAN	O96855	homo sapien	620	7	0.5	693	1	CAUP_DROME	O34269	drosophila
548	7	0.5	590	1	HMAA_DROME	P29555	drosophila	621	7	0.5	697	1	AN3_XENLA	P24316	xenopus lae
549	7	0.5	590	1	NPAL_HUMAN	O99742	homo sapien	622	7	0.5	699	1	KI22_STRPU	P46872	strongyloce
550	7	0.5	592	1	ODP2_DICDI	P36413	dictyosteli	623	7	0.5	699	1	NP14_HUMAN	O14978	homo sapien
551	7	0.5	592	1	XYNB_PSEFL	P23030	pseudomonas	624	7	0.5	699	1	SRCH_HUMAN	P23337	homo sapien
552	7	0.5	593	1	CC23_SCHPO	O42709	schistosach	625	7	0.5	701	1	KF3A_MOUSE	P28177	mus musc
553	7	0.5	594	1	NPAL_MOUSE	P74759	mus musc	626	7	0.5	701	1	UBF1_XENLA	P25960	xenopus lae
554	7	0.5	594	1	RBUX_DROME	P8159	drosophila	627	7	0.5	702	1	DDX4_MOUSE	O61496	mus musc
555	7	0.5	598	1	RACA_DICDI	P34147	dictyosteli	628	7	0.5	702	1	KF3A_HUMAN	O94456	homo sapien
556	7	0.5	600	1	LAM2_CHICK	P14732	gallus gall	629	7	0.5	704	1	CT67_HUMAN	O94453	homo sapien
557	7	0.5	600	1	XKS1_YEAST	P42826	saccharomyc	630	7	0.5	706	1	C1AA_PAEPP	P41777	rattus norv
558	7	0.5	601	1	DEAD_BUCAL	P57453	buchnera ap	631	7	0.5	706	1	CT67_MOUSE	O45338	paenibacilli
559	7	0.5	601	1	DEAD_BUCAL	O83916	buchnera ap	632	7	0.5	706	1	FTSH_MOUSE	P59114	mus musc
560	7	0.5	602	1	DEAD_BUCBP	O89419	buchnera ap	633	7	0.5	706	1	DREB_MOUSE	O9466	mus musc
561	7	0.5	603	1	BUD8_YEAST	P41658	saccharomyc	634	7	0.5	706	1	KIF2_MOUSE	O9466	mus musc
562	7	0.5	604	1	DED1_YEAST	P06634	saccharomyc	635	7	0.5	706	1	KIF2_MOUSE	O9466	mus musc
563	7	0.5	604	1	MTG8_HUMAN	O06455	homo sapien	636	7	0.5	707	1	DREB_RAT	P28743	saccharomyc
564	7	0.5	605	1	WIS1_SCHPO	P33886	schistosach	637	7	0.5	709	1	PP22_YEAST	O07266	rattus norv
565	7	0.5	606	1	MAK7_HUMAN	O43318	homo sapien	638	7	0.5	710	1	GSOD_ERWCH	O01565	erwinia chr
566	7	0.5	610	1	SANI_YEAST	P22470	saccharomyc	639	7	0.5	712	1	GSBP_ERWCH	P31700	erwinia chr
567	7	0.5	611	1	XYNB_PSEFL	P14738	pseudomonas	640	7	0.5	713	1	DDX4_RAT	O64060	rattus norv
568	7	0.5	613	1	DEAD_HA1IN	P44566	haemophilus	641	7	0.5	715	1	ADSV_BOVIN	O28046	bos taurus
569	7	0.5	614	1	DDX5_MOUSE	P17844	homo sapien	642	7	0.5	720	1	NRG3_HUMAN	O01836	caenorhabdi
570	7	0.5	614	1	NRD1_HUMAN	O61656	mus musc	643	7	0.5	720	1	PKP1_BOVIN	P56975	homo sapien
571	7	0.5	614	1	DBP1_YEAST	P20333	homo sapien	644	7	0.5	723	1	MAS2_RH1IV	O93747	rhizobium l
572	7	0.5	618	1	ORC2_DROME	P24784	saccharomyc	645	7	0.5	723	1	MAS2_RH1ME	O92454	rhizobium m
573	7	0.5	618	1	ZEST_DROVI	O24168	drosophila	646	7	0.5	724	1	MDX4_HUMAN	O94453	homo sapien
574	7	0.5	618	1	DNK2_SYNP7	O24762	drosophila	647	7	0.5	727	1	MEPI_ARATH	O91410	arabidopsis
575	7	0.5	621	1	YRT1_CAEEL	P68833	clostridium	648	7	0.5	727	1	VIV_ORYSA	O28161	bos taurus
576	7	0.5	621	1	PPID_BUCAP	O83987	buchnera ap	649	7	0.5	728	1	CDK9_HUMAN	P36195	mus musc
577	7	0.5	622	1	RN12_HUMAN	P40750	baclillus su	650	7	0.5	730	1	GLN3_YEAST	O60505	homo sapien
578	7	0.5	624	1	DEAD_ECOLI	O39492	homo sapien	651	7	0.5	730	1	BAF1_YEAST	P46551	caenorhabdi
579	7	0.5	624	1	DNK2_SYNP7	P23304	escherichia	652	7	0.5	731	1	RRP6_YEAST	P14164	saccharomyc
580	7	0.5	628	1	ARPB_ECOLI	Q9416	homo sapien	653	7	0.5	731	1	DBP1_YEAST	O61955	mus musc
581	7	0.5	632	1	DNK2_SYNP7	P50021	synecococc	654	7	0.5	741	1	KI21_STRPU	P46871	strongyloce
582	7	0.5	634	1	HMPI_CANAL	P46593	candida alb	655	7	0.5	742	1	KM65_YEAST	O03656	saccharomyc
583	7	0.5	634	1	DEAD_KLEPN	O15209	homo sapien	656	7	0.5	742	1	GYP7_YEAST	P48363	saccharomyc
584	7	0.5	634	1	DEAD_KLEPN	O13370	schistosach	657	7	0.5	746	1	KF3B_HUMAN	O15066	homo sapien
585	7	0.5	636	1	DEED_SCHPO	Q01546	homo sapien	658	7	0.5	747	1	ASP4_BOVIN	O61717	mus musc
586	7	0.5	638	1	Y551_SYNY3	P54123	synecocyst	659	7	0.5	747	1	PARC_HA1IN	P43702	haemophilus
587	7	0.5	641	1	PIB2_PENMA	P33573	petromyzon	660	7	0.5	748	1	GUNC_PSEFL	P27033	pseudomonas
588	7	0.5	642	1	DEAD_KLEPN	P23306	klebsiella	661	7	0.5	748	1	KHL1_HUMAN	O94174	mus musc
589	7	0.5	644	1	LONH_METJA	O58812	methanococc	662	7	0.5	751	1	TKR_DROME	P14083	drosophila
590	7	0.5	649	1	DD17_HUMAN	Q43235	homo sapien	663	7	0.5	753	1	ASP4_BOVIN	O28056	bos taurus
591	7	0.5	650	1	Y411_HUMAN	P18302	gallus gall	664	7	0.5	754	1	2ASD_YEAST	P38953	saccharomyc
592	7	0.5	652	1	DREB_CHICK	P12338	saccharomyc	665	7	0.5	757	1	AMY_GLOAB	P23671	clostridium
593	7	0.5	652	1	HS77_YEAST	O68040	rhododactyl	666	7	0.5	760	1	GLH1_CAEEL	P34689	caenorhabdi
594	7	0.5	654	1	ACSA_RHOCA	P53955	saccharomyc	667	7	0.5	763	1	CSW_DROVI	P17480	homo sapien
595	7	0.5	656	1	YNE7_YEAST	P15067	escherichia	668	7	0.5	764	1	UBF1_HUMAN	P25977	rattus norv
596	7	0.5	657	1	GLGX_ECOLI	O15523	homo sapien	669	7	0.5	764	1	UBF1_RAT	P25976	mus musc
597	7	0.5	661	1	PI10_MOUSE	O00571	mus musc	670	7	0.5	765	1	NEFL_MOUSE	O44940	h nuclear f
598	7	0.5	661	1	DDX3_MOUSE	O62167	mus musc	671	7	0.5	772	1	KLPI_HUMAN	P46870	chlamydomon
599	7	0.5	661	1	VASA_DROME	O92167	mus musc	672	7	0.5	776	1	NIMI_NEUCR	P48479	neurospora
600	7	0.5	661	1	MBB1_CHIRE	O9fns4	chlamydomon	673	7	0.5	779	1	KMHC_DICDI	P34125	dictyosteli
601	7	0.5	662	1	YFL8_YEAST	P43620	saccharomyc	674	7	0.5	783	1	SOK2_YEAST	P53438	saccharomyc
602	7	0.5	663	1	PIT_DROME	O9vds1	drosophila	675	7	0.5	785	1	FL10_CHIRE	P46689	chlamydomon
603	7	0.5	663	1	KUOS_YEAST	P47005	saccharomyc	676	7	0.5	789	1	ADOT_RAT	O63180	rattus norv
604	7	0.5	665	1	KE22_HUMAN	Q14807	homo sapien	677	7	0.5	790	1	KIF9_HUMAN	O94453	homo sapien
605	7	0.5	666	1	NOD_DROME	P81805	drosophila	678	7	0.5	790	1	KIF9_MOUSE	O94453	homo sapien
606	7	0.5	666	1	CHS5_YEAST	Q12114	saccharomyc	679	7	0.5	791	1	KEF1_HUMAN	O14782	homo sapien
607	7	0.5	671	1	HSF_KLULA	P2121	kluyveromyc	680	7	0.5	793	1	KEF3C_HUMAN	O92609	homo sapien
608	7	0.5	677	1	UBF1_XENLA	P25979	xenopus lae	681	7	0.5	795	1	TBC5_HUMAN	O35069	mus musc
609	7	0.5	679	1	PAN3_YEAST	P36102	saccharomyc	682	7	0.5	796	1	KF3C_MOUSE	O55165	rattus norv
610	7	0.5	684	1	EP84_HCVVA	P17151	human cytom	683	7	0.5	796	1	KF3C_RAT	P25440	homo sapien
611	7	0.5	684	1	RPFI_HUMAN	P78424	homo sapien	684	7	0.5	801	1	BRD2_HUMAN	O09585	caenorhabdi
612	7	0.5	686	1	PLB3_YEAST	O08108	saccharomyc	685	7	0.5	805	1	DF19_CAEEL		

691	7	0.5	805	1	E2F_DROME	Q27368 drosophila	764	7	0.5	1070	1	AGLU_CANTS	P29064 candida tsu
692	7	0.5	805	1	PIF1_SCHPO	Q9uua2 schizosacch	765	7	0.5	1075	1	FIOS_YEAST	P28894 saccharomyc
693	7	0.5	807	1	YAK1_YEAST	P14680 saccharomyc	766	7	0.5	1077	1	HUES_DROME	Q02308 drosophila
694	7	0.5	808	1	FPEB_DROME	O05192 drosophila	767	7	0.5	1081	1	GALY_YEAST	P19659 saccharomyc
695	7	0.5	810	1	RBB1_YEAST	P21538 saccharomyc	768	7	0.5	1082	1	SP23_YEAST	P35230 saccharomyc
696	7	0.5	816	1	HUNB_DROVI	P13361 drosophila	769	7	0.5	1085	1	AMP1_PLAFO	O66935 plasmodium
697	7	0.5	816	1	YG3A_YEAST	P53278 saccharomyc	770	7	0.5	1085	1	CUT7_SCHPO	P24349 schizosacch
698	7	0.5	817	1	VRP1_YEAST	P37370 saccharomyc	771	7	0.5	1090	1	NIT4_NEUCR	P28349 neurospora
699	7	0.5	829	1	E74A_DROME	P20105 drosophila	772	7	0.5	1093	1	AP17_HUMAN	O55138 homo sapien
700	7	0.5	838	1	L100_ADEP3	O9ytr7 porcine ade	773	7	0.5	1097	1	CCT_DROME	O66433 homo sapien
701	7	0.5	843	1	CO7_FIG	O9tug3 sus petrofa	774	7	0.5	1097	1	KPID_RAT	O35787 rattus norv
702	7	0.5	843	1	CYPI_BRDVA	Q27450 brugia mala	775	7	0.5	1103	1	KPIC_HUMAN	O43856 homo sapien
703	7	0.5	845	1	CLPC_CHLPN	Q948a6 chlamydia p	776	7	0.5	1111	1	KIP1_YEAST	P28742 saccharomyc
704	7	0.5	845	1	ITBX_DROME	P11584 drosophila	777	7	0.5	1115	1	TBC2_CHIRE	Q6vux3 chlamydomon
705	7	0.5	850	1	D7_DICDI	P54682 dictyostella	778	7	0.5	1141	1	SRE2_HUMAN	Q12773 homo sapien
706	7	0.5	850	1	NRG2_HUMAN	O14511 homo sapien	779	7	0.5	1142	1	MGCI_HUMAN	O60732 homo sapien
707	7	0.5	854	1	CC24_YEAST	P11433 saccharomyc	780	7	0.5	1142	1	PAK1_YEAST	P38990 saccharomyc
708	7	0.5	856	1	KF23_HUMAN	Q02241 homo sapien	781	7	0.5	1146	1	YHC3_YEAST	P28742 saccharomyc
709	7	0.5	859	1	ALR1_YEAST	Q08269 saccharomyc	782	7	0.5	1147	1	KIN2_YEAST	P33166 saccharomyc
710	7	0.5	864	1	YC18_HUMAN	Q9u1k2 homo sapien	783	7	0.5	1156	1	PHYB_SOYBN	P24499 glycine max
711	7	0.5	866	1	MYSP_SCHMA	P06198 schistosoma	784	7	0.5	1169	1	YK82_YEAST	P36170 saccharomyc
712	7	0.5	868	1	WV2_MOUSE	O60992 mus musculu	785	7	0.5	1184	1	BIMC_EMENT	P17120 emericella
713	7	0.5	878	1	KPCO_HUMAN	Q96216 homo sapien	786	7	0.5	1185	1	DRPL_HUMAN	P54259 homo sapien
714	7	0.5	878	1	VAV2_HUMAN	P52735 homo sapien	787	7	0.5	1189	1	YJH6_YEAST	P47035 saccharomyc
715	7	0.5	880	1	BRC4_DROME	Q24206 drosophila	788	7	0.5	1202	1	RFM2_YEAST	Q02773 saccharomyc
716	7	0.5	881	1	PRY3_YEAST	P47033 saccharomyc	789	7	0.5	1210	1	AF4_HUMAN	P51825 homo sapien
717	7	0.5	883	1	E74B_DROME	P11536 drosophila	790	7	0.5	1211	1	BUN2_DROME	Q4523 drosophila
718	7	0.5	885	1	YD03_YEAST	O06639 saccharomyc	791	7	0.5	1214	1	BRL1_HUMAN	P55201 homo sapien
719	7	0.5	887	1	K20A_MOUSE	P97329 mus musculu	792	7	0.5	1224	1	ABL1_CAREL	P03949 caenorhabdi
720	7	0.5	890	1	K20A_HUMAN	O95235 homo sapien	793	7	0.5	1227	1	LAF4_HUMAN	P51826 homo sapien
721	7	0.5	892	1	ATX7_HUMAN	O15265 homo sapien	794	7	0.5	1235	1	TRK1_YEAST	P12685 saccharomyc
722	7	0.5	894	1	WPPA_BACSU	P54423 bacillus su	795	7	0.5	1237	1	E75A_DROME	P17671 drosophila
723	7	0.5	902	1	RNE_BUCAL	P57429 bacteriura ap	796	7	0.5	1238	1	YN13_YEAST	P33840 saccharomyc
724	7	0.5	903	1	DPO1_BPR69	O38087 bacterioph	797	7	0.5	1244	1	SLA1_YEAST	P32790 saccharomyc
725	7	0.5	909	1	WTC2_YEAST	P53185 saccharomyc	798	7	0.5	1271	1	Y338_MYCGB	P47580 mycoplasma
726	7	0.5	918	1	PMW1_YEAST	P05030 saccharomyc	799	7	0.5	1298	1	ICP4_HSV11	P08392 herpes slimp
727	7	0.5	925	1	PDC2_YEAST	P30896 saccharomyc	800	7	0.5	1310	1	ICP4_VZVD	P09310 varicella-z
728	7	0.5	928	1	KINH_NEUCR	P48467 neurospora	801	7	0.5	1324	1	SAL1_HUMAN	Q09812 homo sapien
729	7	0.5	928	1	KKRI_YEAST	P34603 saccharomyc	802	7	0.5	1328	1	HUS2_SCHPO	O09811 schizosacch
730	7	0.5	936	1	PHL1_YEAST	P34521 saccharomyc	803	7	0.5	1337	1	YDM5_SCHPO	P08136 schizosacch
731	7	0.5	939	1	ST20_YEAST	Q01497 saccharomyc	804	7	0.5	1341	1	ACTN_HUMAN	O0ukv3 homo sapien
732	7	0.5	939	1	XPC_HUMAN	Q01831 homo sapien	805	7	0.5	1394	1	E75B_DROME	P17672 drosophila
733	7	0.5	946	1	YB16_YEAST	P38250 saccharomyc	806	7	0.5	1396	1	ITAZ_DROME	P12080 drosophila
734	7	0.5	947	1	PMW2_YEAST	P19657 saccharomyc	807	7	0.5	1402	1	SALM_DROVI	P39806 drosophila
735	7	0.5	950	1	URB1_USTMA	P40349 usutlago ma	808	7	0.5	1405	1	Y232_HUMAN	Q02628 homo sapien
736	7	0.5	958	1	MSH4_MOUSE	O99mt2 mus musculu	809	7	0.5	1412	1	ICP4_HSVWG	O02362 marek's dis
737	7	0.5	966	1	FIB1_PETMA	P02674 petromyzon	810	7	0.5	1443	1	B75C_DROME	P13055 drosophila
738	7	0.5	971	1	CLA4_CANAL	O14427 candida alb	811	7	0.5	1444	1	ADP1_MYCGB	P20796 mycoplasma
739	7	0.5	972	1	TOB1_DROME	P30189 drosophila	812	7	0.5	1446	1	IE18_PRYVA	P33479 pseudotabie
740	7	0.5	980	1	KEMS_FELCA	P13369 felis silve	813	7	0.5	1461	1	IE18_PRYVF	P16725 pseudotabie
741	7	0.5	982	1	CRM_DROME	O76906 drosophila	814	7	0.5	1467	1	ICP4_HSVB	P28925 equine hepr
742	7	0.5	989	1	SER4_PLARG	P13823 plasmodium	815	7	0.5	1487	1	ICP4_HSVBK	P17473 equine hepr
743	7	0.5	993	1	RROT_ARATH	O24600 arabidopsis	816	7	0.5	1489	1	VGPO_YEAST	P53115 saccharomyc
744	7	0.5	1004	1	SAU2_MOUSE	Q94x96 mus musculu	817	7	0.5	1509	1	GSRI_HUMAN	Q09zta homo sapien
745	7	0.5	1006	1	EPB6_HUMAN	O15197 homo sapien	818	7	0.5	1531	1	YQ38_CAEEL	Q09459 caenorhabdi
746	7	0.5	1006	1	K125_TOBAC	O23826 nicotiana t	819	7	0.5	1537	1	DNM1_CHICK	Q02072 gallus galli
747	7	0.5	1007	1	SALT2_HUMAN	O9y467 homo sapien	820	7	0.5	1579	1	SSK2_YEAST	P33599 saccharomyc
748	7	0.5	1010	1	WNT5_DROME	P28466 drosophila	821	7	0.5	1581	1	PPRB_HUMAN	O15648 h peroxisom
749	7	0.5	1025	1	KG09_HUMAN	Q910a0 homo sapien	822	7	0.5	1603	1	PSC_DROME	P35880 drosophila
750	7	0.5	1029	1	KSP1_YEAST	P38691 saccharomyc	823	7	0.5	1609	1	FIG2_YEAST	P25653 saccharomyc
751	7	0.5	1029	1	RIP3_RAT	O9ere6 rattus norv	824	7	0.5	1627	1	ADP1_MYCPN	P11311 mycoplasma
752	7	0.5	1036	1	NIT2_NEUCR	P19212 neurospora	825	7	0.5	1659	1	VIT2_FUNHE	Q08293 oncothychnu
753	7	0.5	1047	1	HTRA_DROME	O17468 drosophila	826	7	0.5	1687	1	VIT2_FUNHE	Q12756 homo sapien
754	7	0.5	1051	1	UKL1_MOUSE	O75385 homo sapien	827	7	0.5	1690	1	KFLA_HUMAN	P33173 mus musculu
755	7	0.5	1051	1	UKL1_MOUSE	O70400 mus musculu	828	7	0.5	1695	1	KFLA_MOUSE	P33173 mus musculu
756	7	0.5	1052	1	MGPC_MYCGB	P22747 mycoplasma	829	7	0.5	1719	1	PRD2_HUMAN	Q13029 homo sapien
757	7	0.5	1056	1	K125_ARATH	P82266 arabidopsis	830	7	0.5	1723	1	PM20_CHLPN	Q92812 chlamydia p
758	7	0.5	1057	1	GCSP_PEA	P26969 pium sattiv	831	7	0.5	1772	1	MSPI_PLAFO	P33828 plasmodium
759	7	0.5	1057	1	KFL1_HUMAN	P52732 homo sapien	832	7	0.5	1782	1	VIT2_BOMMO	Q27309 bombyx mori
760	7	0.5	1060	1	SEB4_XENLA	P28025 xenopus lae	833	7	0.5	1807	1	VTAA2_XENLA	P18709 xenopus lae
761	7	0.5	1065	1	SEB4_YEAST	P23365 saccharomyc	834	7	0.5	1816	1	KFLB_HUMAN	O60333 homo sapien
762	7	0.5	1066	1	KL61_DROME	P46863 drosophila	835	7	0.5	1816	1	KFLB_MOUSE	O60575 mus musculu
763	7	0.5	1067	1	BG52_XENLA	Q91783 xenopus lae	836	7	0.5	1822	1	ZAP3_HUMAN	P49750 homo sapien

837	7	0.5	1823	1	VIT_ICHUN	Q91062	ichthyomyzo	910	6	0.5	122	1	MP22_PHLPR	P43214	phileum grat
838	7	0.5	1912	1	Y468_MYCPN	P75109	mycoplasma	911	6	0.5	122	1	PA2_VIRPB	P31854	viperia беру
839	7	0.5	1822	1	VIT1_CHICK	P87409	gallus gall	912	6	0.5	122	1	RL7_STRAM	Q99666	staphylococ
840	7	0.5	1956	1	ATX1_PLAFA	Q04956	plasmodium	913	6	0.5	122	1	YL15_ADE41	P23672	human adeno
841	7	0.5	1960	1	TF20_HUMAN	Q09800	homo sapien	914	6	0.5	124	1	CC26_YEAST	P14724	saccharomyc
842	7	0.5	1983	1	TF20_MOUSE	Q98968	mus musculu	915	6	0.5	124	1	RNP_GAMDR	P00670	camelus dro
843	7	0.5	2004	1	CHDB_HUMAN	Q9HCK8	homo sapien	916	6	0.5	125	1	Y394_RICPR	Q92dd7	ricicetia
844	7	0.5	2038	1	FSH_DROME	P13709	dtrosophila	917	6	0.5	125	1	YB75_METJA	Q58807	methanococc
845	7	0.5	2148	1	VIT1_AEADAE	Q16927	aedes aegypt	918	6	0.5	126	1	YBGS_ECOLI	P75758	escherichia
846	7	0.5	2167	1	BEM2_YEAST	P39660	saccharomyc	919	6	0.5	128	1	Y050_BPT4	P13103	bacterioph
847	7	0.5	2174	1	TF240_HUMAN	Q9uhv7	homo sapien	920	6	0.5	130	1	OA2_DROVI	O44555	dtrosophila
848	7	0.5	2182	1	CAB1_RAT	Q88460	rattus norv	921	6	0.5	130	1	RS8_BUCAP	P59030	buchnera ap
849	7	0.5	2220	1	CAB1_HUMAN	Q9y610	homo sapien	922	6	0.5	131	1	V092_FOWPV	Q91568	fowipox vir
850	7	0.5	2505	1	CCAA_HUMAN	Q00555	homo sapien	923	6	0.5	131	1	YOHF_BACSU	P54514	bacillus su
851	7	0.5	2653	1	CENB_HUMAN	Q02224	homo sapien	924	6	0.5	132	1	RS19_PYRAB	Q9v459	pyrococcus
852	7	0.5	3097	1	CADN_DROME	Q15943	dtrosophila	925	6	0.5	132	1	RS19_PYRFU	Q8u002	pyrococcus
853	7	0.5	3174	1	CHAC_HUMAN	Q96r17	homo sapien	926	6	0.5	132	1	RS19_PYRHO	O59422	pyrococcus
854	7	0.5	3255	1	POLG_LMYO	P31939	1 genome po	927	6	0.5	133	1	RBFA_CHLMU	Q9pku1	chlamydia m
855	7	0.5	3255	1	POLG_LMYO	P89876	1 genome po	928	6	0.5	133	1	YIM5_BPPH1	P10429	bacterioph
856	7	0.5	3969	1	HRX_HUMAN	Q03164	homo sapien	929	6	0.5	134	1	RS24_YEAST	P26782	saccharomyc
857	7	0.5	4705	1	PAT2_DROME	Q9vw71	dtrosophila	930	6	0.5	135	1	YNP0_YEAST	P53902	saccharomyc
858	7	0.5	5147	1	PAT_DROME	P33450	dtrosophila	931	6	0.5	137	1	UCR7_SCHPO	O74533	saccharomyc
859	6	0.5	37	1	PK36_EUGGR	P21532	euglena gra	932	6	0.5	137	1	YEBD_ECOLI	P31063	escherichia
860	6	0.5	40	1	VIT1_MEJGA	P56531	meleagris g	933	6	0.5	138	1	ATPE_WIGBR	Q8d312	wig91eswort
861	6	0.5	50	1	R331_MYCPU	Q98q77	mycoplasma	934	6	0.5	138	1	HEX9_ADE07	P03283	human adeno
862	6	0.5	65	1	V07K_CLV	P28898	carriation 1	935	6	0.5	138	1	IF1A_SCHPO	P55877	schizosacch
863	6	0.5	67	1	Y012_BACAN	Q9rn20	bacillus an	936	6	0.5	139	1	YN06_YEAST	P53842	saccharomyc
864	6	0.5	69	1	RL38_HUMAN	P23411	homo sapien	937	6	0.5	139	1	YVYF_BACSU	P39807	bacillus su
865	6	0.5	69	1	RL38_LYCES	P46291	lycopersico	938	6	0.5	140	1	RLTD_ARATH	Q91217	arabidopsis
866	6	0.5	69	1	RL38_MOUSE	Q9j118	mus musculu	939	6	0.5	140	1	RR8_EUGGR	P21508	euglena gra
867	6	0.5	73	1	RL7_STRAU	P48860	staphylococ	940	6	0.5	140	1	YF62_ARCFU	Q28710	archaeoglob
868	6	0.5	74	1	CH11_STROI	P29115	streptomycc	941	6	0.5	141	1	PER_DROSR	Q04537	dtrosophila
869	6	0.5	77	1	S18W_HUMAN	Q9uh62	homo sapien	942	6	0.5	141	1	RNBR_GIRCA	Q29542	giraffa cam
870	6	0.5	78	1	KCRU_PIG	Q29577	sus scrofa	943	6	0.5	142	1	DIMI_HUMAN	O18433	homo sapien
871	6	0.5	79	1	RS18_UREPA	Q9pfr8	ureaplasma	944	6	0.5	142	1	DIMI_SCHPO	P87215	schizosacch
872	6	0.5	80	1	DBHL_RICCN	Q92h14	ricicetia	945	6	0.5	142	1	Y742_CHLPO	Q92793	chlamydia p
873	6	0.5	80	1	DBHL_RICPR	Q92d26	ricicetia	946	6	0.5	145	1	HA17_CLOBO	P46083	clostridium
874	6	0.5	80	1	IF1_DEIRA	Q9rsk1	delinococcus	947	6	0.5	145	1	YE1A_METJA	P81328	methanococc
875	6	0.5	80	1	YSPF_SHIFL	O55298	shigella fl	948	6	0.5	145	1	YHP2_MYCCA	P45614	mycoplasma
876	6	0.5	88	1	YEAD_SCHPO	O13973	schizosacch	949	6	0.5	146	1	ADP6_ARATH	Q92z62	arabidopsis
877	6	0.5	89	1	SLTB_BRH30	P08027	bacterioph	950	6	0.5	146	1	VG4_SPVIR	P15895	spiroplasma
878	6	0.5	89	1	Y150_ARCFU	Q30087	archaeoglob	951	6	0.5	147	1	SPAM_SALTY	P40612	salmonella
879	6	0.5	89	1	YE96_CLOPE	Q8xk44	clostridium	952	6	0.5	147	1	Y211_MYCGE	P47453	mycoplasma
880	6	0.5	90	1	DBHL_RICRI	Q68451	ricicetia	953	6	0.5	147	1	YD05_METJA	O58701	methanococc
881	6	0.5	90	1	HSEB_RAT	Q9ukf5	rattus norv	954	6	0.5	148	1	GM6_TRYBG	Q26755	trypanosoma
882	6	0.5	90	1	Y060_NPPOP	O10317	oryzia pseu	955	6	0.5	148	1	PTRI_METJA	O57615	methanococc
883	6	0.5	91	1	SMI3_CABEL	P55853	caenorhabd1	956	6	0.5	148	1	YG4E_YEAST	P42337	saccharomyc
884	6	0.5	97	1	IM8A_HUMAN	O60220	homo sapien	957	6	0.5	149	1	BTI1_YEAST	P40314	saccharomyc
885	6	0.5	97	1	IM8A_RAT	Q9wva1	rattus norv	958	6	0.5	149	1	YM66_CLOPE	O8x155	clostridium
886	6	0.5	103	1	KACS_RABIT	P01841	oryctolagus	959	6	0.5	150	1	YHY5_SCHPO	O60154	schizosacch
887	6	0.5	104	1	YJH8_YEAST	P47070	saccharomyc	960	6	0.5	152	1	FLIJ_BUCAP	Q8k441	buchnera ap
888	6	0.5	105	1	V248_FOWPV	P14363	fowipox vir	961	6	0.5	152	1	Y447_WIGBR	O8dca7	wig91eswort
889	6	0.5	105	1	Y613_ARCFU	P29642	archaeoglob	962	6	0.5	153	1	VE6_HPV2A	P25484	human papil
890	6	0.5	106	1	RNT1_TR1HA	P26875	trichoderma	963	6	0.5	154	1	PASC_ECOLI	P45599	escherichia
891	6	0.5	106	1	THIO_GEOCY	Q96952	geodia cydo	964	6	0.5	154	1	PFDS_SCHPO	O94037	schizosacch
892	6	0.5	107	1	FKB1_XENIA	O42123	xenopus lae	965	6	0.5	155	1	ITRF_MAIZE	P01088	zea mays (m
893	6	0.5	107	1	YP94_YEAST	O06835	saccharomyc	966	6	0.5	155	1	PP11_SCHPO	P87051	schizosacch
894	6	0.5	108	1	RL23_MYCGA	O52334	mycoplasma	967	6	0.5	156	1	RNP_MYOGI	Q9xwsl	myoxus glis
895	6	0.5	108	1	SVS4_MOUSE	P18419	mus musculu	968	6	0.5	156	1	YB01_RHOA	O07465	rhodospseudo
896	6	0.5	109	1	PRVA_AMPPE	P02626	amphiuma me	969	6	0.5	157	1	PA1F_HUMAN	P24666	homo sapien
897	6	0.5	111	1	RNPA_MYCPE	O8eu90	mycoplasma	970	6	0.5	157	1	PAIS_HUMAN	P24667	homo sapien
898	6	0.5	111	1	YPJD_BACSU	P42979	bacillus su	971	6	0.5	157	1	PPAC_BOVIN	P11064	bos taurus
899	6	0.5	113	1	RL30_SPOFR	P58375	spodoptera	972	6	0.5	157	1	PPAC_PIG	P81639	sus scrofa
900	6	0.5	114	1	PHS_SUISO	Q97wm6	sulfolobus	973	6	0.5	157	1	YMM3_CABEL	Q23579	caenorhabd1
901	6	0.5	114	1	YOUS_CABEL	P34628	caenorhabd1	974	6	0.5	159	1	PPAC_RAT	P41498	rattus norv
902	6	0.5	115	1	A62F_DROME	O46202	dtrosophila	975	6	0.5	161	1	YP49_RALSO	Q8xvcs	ratstonia s
903	6	0.5	115	1	KV02_RABIT	P01683	oryctolagus	976	6	0.5	161	1	YL30_CLOPE	Q93966	aeropyrum p
904	6	0.5	116	1	TCLA_MOUSE	P56280	mus musculu	977	6	0.5	162	1	RL30_AERPE	O93966	aeropyrum p
905	6	0.5	118	1	YE17_SYNY3	P72731	synchocyst	978	6	0.5	162	1	Y79A_METJA	P81133	methanococc
906	6	0.5	118	1	YE17_SYNY3	Q92363	schizosacch	979	6	0.5	163	1	YA53_METJA	O58553	methanococc
907	6	0.5	121	1	H2B1_TERTH	P08993	tetrahymena	980	6	0.5	164	1	HITI_YEAST	P46573	saccharomyc
908	6	0.5	121	1	H2B2_TERTH	P08994	tetrahymena	981	6	0.5	167	1	HRS1_MOUSE	Q9gcu4	mus musculu
909	6	0.5	121	1	OMP7_STRAU	P21223	staphylococ	982	6	0.5	167	1	ISPF_BUCAP	Q8k9d7	buchnera ap


```

CC -----
DR EMBL; X52482; CA36726.1; -
DR EMBL; D13228; BAA03508.1; -
DR EMBL; X87941; CAA61183.1; -
DR EMBL; Z73018; CAA97261.1; -
DR EMBL; S61041; AAD13922.1; -
DR PIR; S57698; S57698.
DR SGD; S0003465; PHO81.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR004331; SPX.
DR Pfam; PF00023; ANK; 6.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS50088; ANK_REPEAT; 2.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR ANK repeat; Repeat.
KW DOMAIN 198 253 ASN/ASP-RICH.
FT DOMAIN 231 248 POLY-ASN.
FT REPEAT 423 452 ANK 1.
FT REPEAT 458 487 ANK 2.
FT REPEAT 506 535 ANK 3.
FT REPEAT 556 586 ANK 4.
FT REPEAT 591 620 ANK 5.
FT REPEAT 624 653 ANK 6.
FT CONFLICT 248 248 N -> NN (IN REF. 2).
FT CONFLICT 728 728 T -> I (IN REF. 2).
FT CONFLICT 762 762 S -> F (IN REF. 2).
FT CONFLICT 845 845 D -> H (IN REF. 2).
FT CONFLICT 873 873 N -> K (IN REF. 2).
FT CONFLICT 920 920 MISSING (IN REF. 1).
FT CONFLICT 984 984 A -> V (IN REF. 2).
SQ SEQUENCE 1178 AA; 134028 MW; 9314EDB94B3F667D CRC64;

```

```

Query Match 1.3%; Score 17; DB 1; Length 1178;
Best Local Similarity 100.0%; Pred. No. 7.3e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 742 NNNNNNNNNNNNNNNNNI 758
Db 233 NNNNNNNNNNNNNNNNNI 249

```

```

RESULT 5
AAC4_DICDI STANDARD; PRT; 317 AA.
ID AAC4_DICDI
AC P14198;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE AAC-rich mRNA clone PIK330 protein (Fragment).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90066348; PubMed=2511421;
RA Shaw D.R., Richter H., Giorda R., Ohmachi T., Emis H.L.;
RT "Nucleotide sequences of Dictyostelium discoideum developmentally
RT regulated cDNAs rich in (AAC) imply proteins that contain clusters of
RT asparagine, glutamine, or threonine."
RL Mol. Gen. Genet. 218:453-459(1989).
CC -I- DEVELOPMENTAL STAGE: THE CONCENTRATION OF AAC-RICH MRNAs IS LOW
CC -I- IN DORMANT SPORES AND GROWING CELLS, BUT INCREASES DURING
CC SPORE-GERMINATION AND MULTICELLULAR DEVELOPMENT.
CC -I- MICELLANEOUS: SEVERAL PROTEINS DERIVE FROM AAC-RICH MRNA, WHICH,
CC DUE TO A FRAMESHIFT ALSO HAVE ACA AND CAA CODONS AND THUS ARE
CC ASN-, THR- OR GLN-RICH.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC -----
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X16523; CAA34530.1; -
DR PIR; S05356; S05356.
DR DictyDb; DD05008; -.
KW Repeat.
FT DOMAIN 54 73 ASN-RICH.
FT DOMAIN 314 317 ASN-RICH.
FT NON TER 317 317
SQ SEQUENCE 317 AA; 35893 MW; 7FB8B52428D8D2CFC CRC64;

```

```

Query Match 1.2%; Score 16; DB 1; Length 317;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 742 NNNNNNNNNNNNNNNNNI 757
Db 54 NNNNNNNNNNNNNNNNNI 69

```

```

RESULT 6
MYB_DICDI STANDARD; PRT; 451 AA.
ID MYB_DICDI
AC P34127;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE MYB-like protein (Fragment).
GN MYBA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92195700; PubMed=1549373;
RA Stöber-Graesser U., Brydolf B., Bin X., Graesser F., Firtel R.A.,
RA Lipsick J.S.;
RT "The Myb DNA-binding domain is highly conserved in Dictyostelium
RT discoideum."
RL Oncogene 7:589-596(1992).
CC -I- FUNCTION: MAY CONTROL CELLULAR DIFFERENTIATION.
CC -I- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -I- SIMILARITY: Contains 3 myb-like domains.
CC -I- CAUTION: IT IS UNCERTAIN WHETHER MET-1, MET-6, MET-110, MET-114
CC OR MET-123 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z11534; CAB37862.1; -
DR HSSP; P06876; IMBG.
DR DictyDb; DD05044; myba.
DR InterPro; IPR001005; MYB_DNA_binding.
DR Pfam; PF00249; myb_DNA-binding; 3.
DR SMART; SM00717; SANT; 3.
DR PROSITE; PS00037; MYB_1; 3.
DR PROSITE; PS00334; MYB_2; 3.
DR PROSITE; PS50090; MYB_3; 3.
KW Nuclear protein; DNA-binding; Repeat.
FT DNA_BIND 144 195 MYB 1.
FT DNA_BIND 196 247 MYB 2.
FT DNA_BIND 248 298 MYB 3.
FT DOMAIN 25 61 ASN-RICH.
FT DOMAIN 89 139 ASN-RICH.
FT DOMAIN 353 363 PRO-RICH.
FT DOMAIN 48 61 POLY-ASN.
FT DOMAIN 62 67 POLY-GLU.

```



```

FT DOMAIN 68 74 POLY-ASP.
FT DOMAIN 108 132 POLY-ASN.
FT DOMAIN 338 349 POLY-THR.
FT DOMAIN 429 451 POLY-ASN.
FT NON_TER 451 451
SO SEQUENCE 451 AA; 51298 MW; E199B3B471728F60 CRC64;

Query Match 1.2%; Score 16; DB 1; Length 451;
Matches 16; Conservative 100.0%; Pred. No. 2.8e-07; Mismatches 0; Indels 0; Gaps 0;
QY 742 NNNNNNNNNNNNNNNNNN 757
Db 429 NNNNNNNNNNNNNNNNNN 444

RESULT 7
CAR3_DICD1
AC P35352; STANDARD; PRT; 490 AA.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cyclic AMP receptor 3.
OS CARC OR CAR3.
OC Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=93170666; PubMed=8382181;
RA Johnson R.L., Saxe C.L. II, Gollig R., Kimmel A.R., Devreotes P.N.;
RT "Identification and targeted gene disruption of CAR3, a CAMP receptor
RT subtype expressed during multicellular stages of Dictyostelium
RT development."
RL Genes Dev. 7:273-282 (1993).
CC -1- FUNCTION: RECEPTOR FOR CAMP. COORDINATES THE AGGREGATION
CC OF INDIVIDUAL CELLS INTO A MULTICELLULAR ORGANISM AND REGULATES
CC THE EXPRESSION OF A LARGE NUMBER OF DEVELOPMENTALLY REGULATED
CC GENES. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- DEVELOPMENTAL STAGE: INDUCED AT EARLY AGGREGATION (6 HRS) AND
CC MAXIMALLY EXPRESSED AT THE WOUND STAGE (9-12HRS), LEVEL OF
CC EXPRESSION PEAKS AGAIN DURING THE SLUG STAGE (18HRS) AND DECLINES
CC AT CUMINATION.
CC -1- PTM: CARBOXY-TERMINAL SER OR THR RESIDUES MAY BE PHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO FAMILY 5 OF G-PROTEIN COUPLED RECEPTORS.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-15 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S55235; AAB25437.1; -.
DR Dictydb; DR02033; GARC.
DR InterPro; IPR000848; GPCR_CAM.
DR InterPro; IPR000832; GPCR_secretin.
DR PRINTS; PR00247; GPCRCAMP.
DR PROSITE; PS50261; G_PROTEIN_RECPE_F2_4; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Multigene family.
FT DOMAIN 1 23
FT TRANSSEM 1 23 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 24 43 1 (POTENTIAL).
FT TRANSSEM 44 57 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 58 78 2 (POTENTIAL).
FT TRANSSEM 79 94 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 95 120 3 (POTENTIAL).
FT DOMAIN 121 131 CYTOPLASMIC (POTENTIAL).

```

Query Match	Best Local Similarity	Score 16;	DB 1;	Length 490;
Matches 16;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Db	742	399	NNNNNNNNNNNNNNNNNN	757
SEQUENCE	490 AA;	56161 MM;	A28BA834D8626153	CRC64;
FT TRANSMEM	132	150	4 (POTENTIAL).	
FT DOMAIN	151	173	EXTRACELLULAR (POTENTIAL).	
FT TRANSMEM	174	192	5 (POTENTIAL).	
FT DOMAIN	193	216	CYTOPLASMIC (POTENTIAL).	
FT TRANSMEM	217	235	6 (POTENTIAL).	
FT DOMAIN	236	246	EXTRACELLULAR (POTENTIAL).	
FT TRANSMEM	247	271	7 (POTENTIAL).	
FT DOMAIN	272	490	CYTOPLASMIC (POTENTIAL).	
FT MOD_RES	204	204	PHOSPHORYLATION (BY PKA) (POTENTIAL).	
FT DOMAIN	303	439	ASN-RICH.	
FT DOMAIN	399	427	POLY-ASN.	
SEQUENCE	490 AA;	56161 MM;	A28BA834D8626153	CRC64;
Query Match	Best Local Similarity	Score 16;	DB 1;	Length 490;
Matches 16;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Db	742	399	NNNNNNNNNNNNNNNNNN	757
SEQUENCE	490 AA;	56161 MM;	A28BA834D8626153	CRC64;


```

DR Pfam; PF00096; zf-C2H2; 2.
DR SMART; SM00355; ZNF_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW Nuclear protein; Zinc-finger; Metal-binding; Repeat.
FT ZN_FING 346 368 C2H2-TYPE 1.
FT ZN_FING 374 397 C2H2-TYPE 2.
FT DOMAIN 8 35 POLY-GLN.
FT DOMAIN 98 104 POLY-ASN.
FT DOMAIN 143 157 POLY-ASN.
FT DOMAIN 173 177 POLY-ALA.
FT DOMAIN 240 245 POLY-HIS.
FT DOMAIN 417 420 POLY-SER.
FT DOMAIN 421 433 POLY-ASN.
FT DOMAIN 441 450 POLY-ALA.
SQ SEQUENCE 490 AA; 54382 MW; 4D0DA8DE43F171ED CRC64;

Query Match 1.2%; Score 16; DB 1; Length 490;
Best Local Similarity 100.0%; Pred. No. 3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 743 NNNNNNNNNNNNNNNNNI 758
Db 143 NNNNNNNNNNNNNNNNNI 158

RESULT 9
HSF1 ARATH STANDARD; PRT; 495 AA.
ID HSF1 ARATH STANDARD; PRT; 495 AA.
AC P4151; O23615;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Heat shock factor protein 1 (HSF 1) (Heat shock transcription factor
DE 1) (HSF 1).
GN HSF1 OR AT4G1750 OR DL4910C.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=95036006; PubMed=7948881;
RA Huebel A., Schoeffl F.;
RT "Arabidopsis heat shock factor: isolation and characterization of the
RT gene and the recombinant protein.";
RL Plant Mol. Biol. 26:353-362(1994).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=9612113; PubMed=9461215;
RA Bevan M., Bancroft I., Bent E., Love K., Goodman H.M., Dean C.,
RA Bergkamp R., Dirkse W., van Staveren M., Stiekema W., Drost L.,
RA Ridley P., Hudson S.-A., Patel K., Murphy G., Piffeneill P.,
RA Weiler H., Weiler E., Wambutt R., Weitzenecker T., Pohl T., Terry N.,
RA Gielen J., Villarroel R., De Clercq R., van Montagu M., Lecharny A.,
RA Aubourg S., Gylis I., Kreis M., Lao N., Kavanagh T., Hempel S.,
RA Kotter P., Entian K.-D., Rieger M., Schaefer M., Funk B.,
RA Puigdomenech P., Douka A., Vouklatou E., Milioni D., Hatzopoulos P.,
RA Pirvandi E., Obermayer B., Hilbert H., Duesterhoeft A., Moores T.,
RA Jones J.D.G., Eneva T., Palme K., Benes V., Rechmann S., Anserge W.,
RA Cooke R., Berger C., Delseny M., Voet M., Volckaert G., Mewes H.-W.,
RA Kloterman S., Schueller C., Chalvatzis N.;
RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of
RT Arabidopsis thaliana.";
RL Nature 391:485-488(1998).
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,

```

```

RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,
RA Harris B., Anserge W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermayer B., Maché R., Mueller M.,
RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidthini T.,
RA Reichert B., Potteville D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hohenseil J., Zimmermann W., Weiler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schuren J., Grymoprez B., Chuang Y.-J., Vandebussche F.,
RA Braeken M., Welfens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weitzenecker T., Bothe G., Rampsperger U., Hilbert H., Braun M.,
RA Holzner E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Moollman P., Klein Lanhorst R., Rose M., Haut J., Koetter P.,
RA Berneriser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA de Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Leonard N., McIay K., Mayes R.,
RA Petter A., Rajendram M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Grandjean K., Dauner D., Herzl A.,
RA Neumann S., Argilou A., Vitale D., Liguori R., Pirvandi E.,
RA Maesenet O., Outgley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chedof F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Bargas M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., France P., Bietke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang B., Spiegel L.,
RA Sehon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courneye J., Cloud J., Abbott A., Scott K., Johnson D.,
RA Mix P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pezin K., Hillier L.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antoniou B., Zifanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong Y., Preston R., Vil D., Shekhar M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
RA Grnec S., Shohdy N., Haasegawa A., Hamed A., Lohdi M., Johnson A.,
RA Chen E., Marra M., Martensen R., Mccombe W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:769-777(1999).
CC -!- FUNCTION: DNA-BINDING PROTEIN THAT SPECIFICALLY BINDS HEAT SHOCK
CC PROMOTER ELEMENTS (HSE) AND ACTIVATES TRANSCRIPTION (BY
CC SIMILARITY).
CC -!- SUBUNIT: Homotrimer.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- PTM: EXHIBITS TEMPERATURE-DEPENDENT PHOSPHORYLATION
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE HSF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annouce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X76167; CA53761.1; -.
DR EMBL; Z97344; CAB10555.1; -.
DR EMBL; AL161547; CAB8778.1; -.
DR PIR; F71447; S52641.
DR HSSP; P22121; 3HSF.
DR TRANSFAC; T04394; -.
DR InterPro; IPR000232; HSF_DNA_bind.
DR InterPro; IPR002341; HSF_ETS.
DR Pfam; PF00447; HSF_DNA_bind; 1.
DR PRINTS; PR00056; HSFDOMAIN.
DR ProDom; PD001788; HSF_DNA_bind; 1.

```



```

OY 742 NNNNNNNNNNNNNNNN 757
Db 419 NNNNNNNNNNNNNNNN 434

RESULT 11
KMBB D1CD1
ID KMBB D1CD1 STANDARD; PRT; 732 AA.
AC P90648;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myosin heavy chain kinase B (BC 2.7.1.129) (MHCK B).
GN MHCKB OR MHCKB.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
NCBI_TaxID=44689;
[1]
SEQUENCE FROM N.A.
RP STRAIN=AX3;
RC MEDLINE=97277316; PubMed=9115238;
RX Clancy C.E., Mendoza M.G., Natsmith T.V., Kolman M.F., Egelhoff T.T.;
"Identification of a protein kinase from Dictyostelium with homology
to the novel catalytic domain of myosin heavy chain kinase A.";
RL J. Biol. Chem. 272:11812-11815(1997)
CC -1- FUNCTION: PHOSPHORYLATES THREONINE IN THE C-TERMINAL TAIL REGION
OF MYOSIN II HEAVY CHAIN. THIS PHOSPHORYLATION IS CRITICAL IN
REGULATING THE ASSEMBLY AND DISASSEMBLY OF MYOSIN II FILAMENT.
CC -1- CATALYTIC ACTIVITY: ATP + [myosin heavy-chain] = ADP + [myosin
heavy-chain] phosphate.
CC -1- DOMAIN: CONSISTS OF AN N-TERMINAL DOMAIN, A CENTRAL NONREPEATITIVE
CATALYTIC DOMAIN, AND A C-TERMINAL DOMAIN WITH SEVEN WD REPEATS.
CC -1- SIMILARITY: BELONGS TO THE MHCK / EF-2 PROTEIN KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on
its use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U90946; AAB50136.1; -.
DR Dictydb; DD01087; mhkb.
DR InterPro; IPR004166; MHCK_EF2_kinase.
DR InterPro; IPR001680; WD40-.
DR Pfam; PF02816; Alpha_kinase; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINRPT.
DR PRODOM; PD000018; WD40; 2.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00678; WD_REPEATS_1; 5.
DR PROSITE; PS50082; WD_REPEATS_2; 6.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat;
WD Repeat.
KW NP BIND
KW NP BIND 298 303 ATP (POTENTIAL).
KW DOMAIN 355 381 POLY-ASN.
FT REPEAT 458 486 WD 1.
FT REPEAT 500 528 WD 2.
FT REPEAT 540 568 WD 3.
FT REPEAT 580 608 WD 4.
FT REPEAT 620 648 WD 5.
FT REPEAT 660 688 WD 6.
FT REPEAT 700 730 WD 7.
SQ SEQUENCE 732 AA; 83167 MW; A7233C4BD56F4088 CRC64;

Query Match 1.2%; Score 16; DB 1; Length 732;
Best Local Similarity 100.0%; Pred. No. 4,4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


Db 355 NNNNNNNNNNNNNNNNN 370

```

RESULT 12
CIGB DICI
ID_CIGB DICI STANDARD; PRT; 735 AA.
AC 094481;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein cigb (Fragment).
CC
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Eukaryota; Dictyostelidae; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Loomis W.F.;
RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
CC
CC -1- SIMILARITY: TO D.DISCOTHEUM CIGB.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U66528; AAB06791.1; -.
DR DictyDb; DD01058; cigb.
DR InterPro; IPR00315; Znf_Box.
DR Pfam; PF00643; zf-B_box; 1.
KW Repeat.
FT NON_TER 1 1
FT DOMAIN 151 155 POLY-ASN.
FT DOMAIN 251 268 POLY-ASN.
FT DOMAIN 328 667 TANDDEM REPEATS.
SQ SEQUENCE 735 AA; 83443 MW; 9765013D88DA6C8C CRC64;

Query Match 1.2%; Score 16; DB 1; Length 735;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 742 NNNNNNNNNNNNNNNNN 757
DB 253 NNNNNNNNNNNNNNNNN 268

RESULT 13
Y38 YEAST STANDARD; PRT; 758 AA.
ID Y38 YEAST STANDARD; PRT; 758 AA.
AC 003825;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 85.0 kDa protein in HLu1-SMP2 intergenic region.
GN YMR164C OR YMR520.13C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor K., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagsi R., Lyne G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome

```

```

RT XIII."
RL Nature 387:90-93 (1997).
CC
CC -1- SIMILARITY: Contains 1 Lish domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z49705; CAA89800.1; -.
DR PIR; S54522; S54522.
DR SGD; S0004774; MSS11.
DR GO; GO:0005634; C:nucleus; IC.
DR GO; GO:0003704; F:specific RNA polymerase II transcription fa...; IDA.
DR GO; GO:0045944; P:positive regulation of transcription from P...; IDA.
DR GO; GO:0007124; P:pseudophal growth; IGI.
DR GO; GO:0005983; P:starch catabolism; IMP.
DR InterPro; IPR006594; Lish.
DR SMART; SM00667; Lish; 1.
DR PROSITE; PS50896; Lish; 1.
KW Hypothetical protein.
FT DOMAIN 51 83 LISH.
FT DOMAIN 290 329 POLY-GLN.
FT DOMAIN 605 637 POLY-ASN.
FT DOMAIN 653 656 POLY-SER.
SQ SEQUENCE 758 AA; 85050 MW; BA05BFC754D9294B CRC64;

Query Match 1.2%; Score 16; DB 1; Length 758;
Best Local Similarity 100.0%; Pred. No. 4.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 742 NNNNNNNNNNNNNNNNN 757
DB 605 NNNNNNNNNNNNNNNNN 620

RESULT 14
GPR1 YEAST STANDARD; PRT; 961 AA.
ID GPR1 YEAST STANDARD; PRT; 961 AA.
AC 012361;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE G protein-coupled receptor GPR1.
GN GPR1 OR YDL035C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=97197972; PubMed=9046088;
RA Saren A.M., Laamanen P., Lejarregui J.B., Paulin L.;
RT "The sequence of a 36.7 kb segment on the left arm of chromosome IV
RT from Saccharomyces cerevisiae reveals 20 non-overlapping open reading
RT frames (ORFs) including SIR4, FAD1, NAM1, RIM1, SIR2, NAT1, PRP9, ACT2
RT and MPT1 and 11 new ORFs."
RL Yeast 13:65-71 (1997).
CC
CC -1- FUNCTION: Seems to associate with GPA2 and act as G protein-
CC coupled receptor that senses glucose and controls filamentous
CC growth. It acts upstream of adenylate cyclase and is required for

```



```
DR InterPro: IPR000387; TYR phosphatase.
DR InterPro: IPR000242; Tyr_PP.
DR Pfam; PF00102; Tyrosinphatase; 1.
DR PRINTS; PR00700; PR07PHPTASE.
DR SMART; SM00194; PTpc; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00566; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00555; TYR_PHOSPHATASE_PTP; 1.
DR Hydrolase.
KM ACT_SITE 649 649
FT DOMAIN 460 716 BY SIMILARITY.
FT DOMAIN 64 71 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 109 118 POLY-ASN.
FT DOMAIN 137 190 POLY-ASN.
FT DOMAIN 249 257 POLY-SER.
FT DOMAIN 258 265 POLY-THR.
FT DOMAIN 286 289 POLY-ASN.
FT DOMAIN 366 371 POLY-SER.
FT DOMAIN 787 790 POLY-GLN.
FT DOMAIN 834 839 POLY-GLN.
FT DOMAIN 883 892 POLY-GLN.
FT DOMAIN 906 914 POLY-ASN.
FT DOMAIN 943 963 POLY-ASN.
SQ SEQUENCE 989 AA; 109995 MW; 9371105AF80974AF CRC64;

Query Match 1.2%; Score 16; DB 1; Length 989;
Best Local Similarity 100.0%; Pred. No. 5.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 742 NNNNNNNNNNNNNNNNN 757
Db 137 NNNNNNNNNNNNNNNNN 152

RESULT 17
SUZ2 DROME STANDARD; PRT; 1365 AA.
AC P25172;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Suppressor 2 of zeste protein (Protein posterior sex combs).
GN SU(2)2.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=91279476; PubMed=2057369;
RA Brunk B.P., Adler P.N.;
RT "The sequence of the Drosophila regulatory gene Suppressor two of
RT zeste."
RL Nucleic Acids Res. 19:3149-3149(1991).
CC -!- FUNCTION: REGULATES EXPRESSION OF THE HOMEOTIC SELECTOR GENES BY
CC INTERLUENCING HIGHER-ORDER CHROMATIN STRUCTURE THROUGH INTERACTION
CC WITH OTHER PROTEINS.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56798; CAA40134.1; -
CC EMBL; X56799; CAA40135.1; -
CC PIR; S14871; S14871.
```

```
DR FlyBase; FBgn0008654; Su(z)2.
DR InterPro: IPR001841; ZnF_ring.
DR Pfam; PF00097; zf-C3HC4_1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KM Zinc-finger; Developmental protein; DNA-binding; Nuclear protein.
FT ZN_FING 35 74
FT DOMAIN 623 628 POLY-GLN.
FT DOMAIN 1077 1096 POLY-ASN.
FT DOMAIN 1241 1251 POLY-SER.
FT DOMAIN 603 603 MISSING (IN REF. 1; CAA40134).
FT CONFLICT 785 785 K -> N (IN REF. 1; CAA40134).
FT CONFLICT 831 831 A -> R (IN REF. 1; CAA40134).
FT CONFLICT 965 965 MISSING (IN REF. 1; CAA40134).
FT CONFLICT 1065 1065 D -> E (IN REF. 1; CAA40134).
FT CONFLICT 1076 1076 MISSING (IN REF. 1; CAA40134).
FT CONFLICT 1287 1287 A -> P (IN REF. 1; CAA40134).
SQ SEQUENCE 1365 AA; 146058 MW; 7B4BA0F35B0FA683 CRC64;

Query Match 1.2%; Score 16; DB 1; Length 1365;
Best Local Similarity 100.0%; Pred. No. 7.9e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 742 NNNNNNNNNNNNNNNNN 757
Db 1077 NNNNNNNNNNNNNNNNN 1092

RESULT 18
KYK1 DICD1 STANDARD; PRT; 1584 AA.
AC P18160;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Non-receptor tyrosine kinase spore lysis A (EC 2.7.1.112) (Tyrosine-
DE protein kinase 1).
GN PYKA OR SPLA OR DPYK1.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JH10;
RX MEDLINE=97053827; PubMed=8898241;
RA Nuckolls G.H., Oshero N., Loomis W.F., Spudich J.A.;
RT "The Dictyostelium dual-specificity kinase sp1A is essential for
RT spore differentiation."
RL Development 122:3295-3305(1996).
RN [2]
RP SEQUENCE OF 1248-1584 FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in
RT Dictyostelium discoideum."
RL Mol. Cell. Biol. 10:3578-3583(1990).
CC -!- FUNCTION: ESSENTIAL FOR SPORE DIFFERENTIATION.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH A PEAK
CC DURING THE MOUND STAGE OF MORPHOGENESIS.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```


Qy	Db	Query Match	Best Local Similarity	Score 16	DB 1	Length 1858	Matches 16	Conservative	Mismatches	Indels	Gaps
742	185	NNNNNNNNNNNNNNNN 757		1.2%	100.0%	1e-06	0	0	0	0	0
<p>RESULT 22</p> <p>TAGB_DICDI STANDARD; PRT, 1905 AA.</p> <p>ID TAGB_DICDI STANDARD; PRT, 1905 AA.</p> <p>AC P54683;</p> <p>DT 01-OCT-1996 (Rel. 34, Created)</p> <p>DT 01-OCT-1996 (Rel. 34, Last sequence update)</p> <p>DT 28-FEB-2003 (Rel. 41, Last annotation update)</p> <p>DE Prestalk-specific protein tagb precursor (EC 3.4.21.-).</p> <p>DE TAGB.</p> <p>OS Dictyostelium discoideum (slime mold).</p> <p>OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.</p> <p>OX NCBI_TaxID=44689;</p> <p>RN [1]</p> <p>RP SEQUENCE FROM N.A.</p> <p>RC STRAIN=AX4;</p> <p>RX MEDLINE=95262903; PubMed=7744252;</p> <p>RA Shaulsky G., Kuspa A., Loomis W.F.;</p> <p>RT "A multidrug resistance transporter/serine protease gene is required for prestalk specialization in Dictyostelium.";</p> <p>RL Genes Dev. 9:1111-1122(1995).</p> <p>CC -1- FUNCTION: INTERCELLULAR COMMUNICATION VIA TAGB MAY MEDIATE INTEGRATION OF CELLULAR DIFFERENTIATION WITH MORPHOGENESIS.</p> <p>CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO PEPTIDASE FAMILY S8.</p> <p>CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.</p> <p>CC -1- SIMILARITY: STRONG, TO TAGC.</p> <p>CC -----</p> <p>CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - between the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/).</p> <p>CC or send an email to license@isb-sib.ch.</p> <p>CC -----</p> <p>DR EMBL, U20432, AAA62212.1; --</p> <p>DR PIR, T18267, T18267.</p> <p>DR MEROPS, S08, UPW; --</p> <p>DR DictyDB, DD02059, tagB.</p> <p>DR InterPro, IPR003593, AAA_ATPase.</p> <p>DR InterPro, IPR001140, ABC_TM_transp.</p> <p>DR InterPro, IPR003439, ABC_transporter.</p> <p>DR InterPro, IPR000209, peptidase_S8.</p>											

DR	Pfam; PF00664; ABC membrane; 1.
DR	Pfam; PF00005; ABC_tran; 1.
DR	Pfam; PF00082; Peptidase_S8; 1.
DR	PRINTS; PR07723; SUBTILISIN.
DR	ProDom; PD000006; ABC_transporter; 1.
DR	SMART; SM00382; Aaa; 1.
DR	PROSITE; PS00121; ABC_TRANSPORTER_1; 1.
DR	PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
DR	PROSITE; PS00136; SUBTILASE_ASP; FALSE_NEG.
DR	PROSITE; PS00137; SUBTILASE_HIS; 1.
KW	Hydrolase; Serine protease; ATP-binding; Transport; Transmembrane;
KW	Signal.
FT	SIGNAL 1 31 POTENTIAL.
FT	CHAIN 32 1905 PRESTALK-SPECIFIC PROTEIN TAGB.
FT	DOMAIN 378 700 PROTEASE.
FT	DOMAIN 1518 1756 ABC_TRANSPORTER.
FT	TRANSMEM 1011 1031 POTENTIAL.
FT	TRANSMEM 1076 1096 POTENTIAL.
FT	TRANSMEM 1121 1141 POTENTIAL.
FT	TRANSMEM 1210 1230 POTENTIAL.
FT	TRANSMEM 1309 1329 POTENTIAL.
FT	TRANSMEM 1332 1352 POTENTIAL.
FT	TRANSMEM 387 387 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 432 432 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 695 695 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	NP_BIND 1553 1560 ATP (POTENTIAL).
FT	DOMAIN 63 67 POLY-GLN.
FT	DOMAIN 95 104 POLY-ASN.
FT	DOMAIN 107 134 POLY-ASN.
FT	DOMAIN 311 321 POLY-SER.
FT	DOMAIN 833 837 POLY-SER.
FT	DOMAIN 838 844 POLY-GLY.
FT	DOMAIN 871 876 POLY-LEU.
FT	DOMAIN 1012 1015 POLY-TLE.
FT	DOMAIN 1386 1389 POLY-GLU.
FT	DOMAIN 1398 1404 POLY-GLY.
FT	DOMAIN 1445 1450 POLY-ASN.
FT	DOMAIN 1765 1779 POLY-ASN.
FT	DOMAIN 1782 1785 POLY-SER.
FT	DOMAIN 1807 1812 POLY-PRO.
FT	DOMAIN 1813 1860 POLY-GLN.
FT	DOMAIN 1872 1878 POLY-PRO.
FT	CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 747 747 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 823 823 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 1172 1172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 1522 1522 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 1658 1658 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE 1905 AA; 212518 MW; B88223FPA89A8E13C CRC64;
Query Match 1.2%; Score 16; DB 1; Length 1905;	
Best Local Similarity 100.0%; Pred. No. 1.le-06;	
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Oy	742 757
Db	107 NNNNNNNNNNNNNNNNN 122
RESULT 23	
ARP2_PLAFA STANDARD; PRT; 451 AA.	
AC	P13824;
DT	01-JAN-1990 (Rel. 13, Created)
DT	01-JAN-1990 (Rel. 13, Last sequence update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)
DE	Clustered-asparagine-rich protein (Fragment).
OS	Plasmodium falciparum.
OC	Eukaryota; Alveolata; Apicomplexa; Hemosporidia; Plasmodium.
NCBI	TaxID=5833;

RA Ouellette B.F., Clark M.W., Keng T., Storms R.K., Zhong W.W.,
 RA Zeng B., Fortin N., Delaney S., Barton A.B., Kaback D.B., Bussey H.,
 RA "Sequencing of chromosome I from *Saccharomyces cerevisiae*: analysis
 RT of a 32 kb region between the LTR1 and SP07 genes.",
 RL Genome 36:32-42(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94193531; PubMed=8144453;
 RA Barton A.B., Kaback D.B.;
 RT "Molecular cloning of chromosome I DNA from *Saccharomyces cerevisiae*:
 RT analysis of the genes in the FUN38-MAK16-SP07 region.",
 RL J. Bacteriol. 176:1872-1880(1994).
 CC -1- FUNCTION: CONTROL OF ADH GENE EXPRESSION. IT IS REQUIRED FOR THE
 CC EXPRESSION OF GENES INVOLVED IN NONFERMENTATIVE GROWTH AND IT
 CC MEDIATES OR IS REQUIRED FOR THE ACTION OF THE SPT6 & SPT10 GENES.
 CC -1- MISCELLANEOUS: THE 169 CARBOXYL-TERMINAL RESIDUES ARE IMPORTANT
 CC FOR CCR4 FUNCTION.
 CC -1- SIMILARITY: BELONGS TO THE CCR4/NOCTURIN FAMILY.
 CC -1- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; S50459; AAB24455.1; -;
 DR EMBL; L05146; AAC04936.1; -;
 DR PIR; S36713; S36713.
 DR TRANSFAC; T03228; -;
 DR SGD; S0000019; CCR4.
 DR GO; GO:0030015; C:CCR4-NOT core complex; IPI.
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0000175; F:3'-5' exoribonuclease activity; IDA.
 DR GO; GO:0000289; P:poly(A) tail shortening; IDA.
 DR GO; GO:0006557; P:regulation of transcription from Pol II pro. .; IPI.
 DR InterPro; IPR005135; Exo_endo_phos.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR003591; LRR_cyp.
 DR Pfam; PF03372; Exo_endo_phos; 1.
 DR Pfam; PF00560; LRR; 3.
 DR PRINTS; PRO0019; LEURICHRPT.
 DR SMART; SM00369; LRR_TYP; 1.
 DR KM Transcription regulation; Repressor; Leucine-rich repeat; Repeat.
 FT DOMAIN 15 24
 FT POLY-GLN.
 FT 77 103
 FT ASN-RICH.
 FT DOMAIN 89 103
 FT POLY-ASN.
 FT 190 206
 FT GLN-RICH.
 FT REPEAT 356 379
 FT LRR 1.
 FT 380 402
 FT LRR 2.
 FT REPEAT 403 418
 FT LRR 3.
 FT 419 444
 FT LRR 4.
 FT REPEAT 445 467
 FT LRR 5.
 FT 544 544
 FT L -> I (IN REF. 2).
 FT CONFLICT 544 544
 FT E -> V (IN REF. 2).
 FT 803 803
 FT SEQUENCE 837 AA; 94699 MW; B8DCFF24F5A35A9 CRC64;
 SO
 Query Match 1.2%; Score 15; DB 1; Length 837;
 Best Local Similarity 100.0%; Pred. No. 4.7e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DT 01-OCT-1996 (Rel. 34, last sequence update)
 DT 15-SEP-2003 (Rel. 42, last annotation update)
 DE Hypothetical regulatory protein in PDR6-PDR1 intergenic region.
 GN YGL014W OR YGL023.
 OS *Saccharomyces cerevisiae* (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC *Saccharomyces cerevisiae*; Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL125-2B;
 RA MEDLINE=91353083; PubMed=1909073;
 RX Chen W., Balzi E., Capieaux E., Goffeau A.;
 RT "The YGL023 gene encodes a putative regulatory protein.",
 RL Yeast 7:309-312(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL125-2B;
 RA MEDLINE=91353083; PubMed=1882553;
 RX Chen W., Balzi E., Capieaux E., Choder M., Goffeau A.;
 RT "The DNA sequencing of the 17 kb HindIII fragment spanning the LTR1
 RT and ATR1 loci on chromosome VII from *Saccharomyces cerevisiae* reveals
 RT pleiotropic drug resistance.",
 RL Yeast 7:287-299(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Hedling U., Hofmann B., Delius H.;
 RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Is not essential for haploid growth, but may affect
 CC diploid formation.
 CC -1- SIMILARITY: BELONGS TO THE PUMILIO/MPFS FAMILY.
 CC -1- SIMILARITY: SOME, TO YEAST YUR091C.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; S57889; AAB19616.1; -;
 DR EMBL; S58126; AAD13898.1; -;
 DR EMBL; Z72536; CAA86714.1; -;
 DR PIR; S64016; S64016.
 DR SGD; S0002982; PUF4.
 DR GO; GO:0007569; P:cell aging; IGI.
 DR GO; GO:0000288; P:mRNA catabolism, deadenylation-dependent; IGI.
 DR GO; GO:0006005; P:protein targeting; IMP.
 DR InterPro; IPR001313; Pumilio/Puf.
 DR Pfam; PF00806; PUF; 8.
 DR SMART; SM00025; Pumilio; 8.
 DR KM Hypochemical protein; Transcription regulation; Zinc-finger;
 DR Metal-binding; Repeat.
 FT DOMAIN 8 38
 FT ASP/GLU-RICH (ACIDIC).
 FT 471 542
 FT ASN-RICH.
 FT 521 535
 FT POLY-ASN (POTENTIAL ACTIVATING DOMAIN).
 FT ZN FING 724 742
 FT CA-TYPE.
 FT REPEAT 571 583
 FT 583 583
 FT REPEAT 607 619
 FT 643 655
 FT REPEAT 679 691
 FT 716 728
 FT REPEAT 752 764
 FT 831 843
 FT REPEAT 843 843
 FT CONFLICT 595 595
 FT SEQUENCE 888 AA; 97797 MW; 659BA1062439F642 CRC64;
 SO
 Query Match 1.2%; Score 15; DB 1; Length 888;
 Best Local Similarity 100.0%; Pred. No. 5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

Oy 742 NNNNNNNNNNNNNN 756
Db 521 NNNNNNNNNNNNNN 535

RESULT 27
FABI_YEAST STANDARD; PRT; 2278 AA.
AC P34755;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 1-phosphatidylinositol-3-phosphate 5-kinase FABI (EC 2.7.1.150)
DE [phosphatidylinositol-3-phosphate 5-kinase] (Type III PIP kinase).
GN FABI OR YFR019W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95352039; PubMed=7663021;
RA Yamamoto A., Dewald D.B., Boronkov I.V., Anderson R.A., Emr S.D.,
RA Koshland D.;
RT "Novel PI(4)P 5-kinase homologue, Fab1p, essential for normal vacuole
RT function and morphology in yeast.";
RL Mol. Biol. Cell 6:525-539(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Saenuma S.-I., Saenuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae.";
RL Nat. Genet. 10:261-268(1995).
RN [3]
RP CHARACTERIZATION
RX MEDLINE=99035548; PubMed=9811604;
RA Cooke F.T., Dove S.R., McEwen R.K., Painter G., Holmes A.B.,
RA Hall M.N., Mitchell R.H., Parker P.J.;
RT "The stress-activated phosphatidylinositol 3-phosphate 5-kinase Fab1p
RT is essential for vacuole function in S. cerevisiae.";
RL Curr. Biol. 8:1219-1222(1998).
RN [4]
RP SIMILARITY TO TCP1/CCT FAMILY.
RX MEDLINE=95314774; PubMed=7794526;
RA Waldmann T., Lupas A., Kellermann J., Peters J., Baumeister W.;
RT "Primary structure of the thermosome from Thermoplasma acidophilum.";
RL Biol. Chem. Hoppe-Seyler 376:119-126(1995).
CC -1- FUNCTION: Catalyzes the phosphorylation of phosphatidylinositol-3-
CC phosphate on the fifth hydroxyl of the myo-inositol ring, to form
CC phosphatidylinositol-4,5-bisphosphate. Required for endocytic-
CC vacuolar pathway and nuclear migration. The product of the
CC reaction it catalyzes functions as an important regulator of
CC vacuole homeostasis perhaps by controlling membrane flux to and/or
CC from the vacuole.
CC -1- CATALYTIC ACTIVITY: ATP + 1-phosphatidy1-D-myo-inositol 4-
CC monophosphate = ADP + 1-phosphatidy1-D-myo-inositol 4,5-
CC bisphosphate.
CC -1- COFACTOR: Magnesium or Manganese.
CC -1- SUBCELLULAR LOCATION: VACUOLE-ASSOCIATED (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE PTING(4)P-5-KINASE FAMILY.
CC -1- SIMILARITY: Contains 1 FYVE-type zinc finger.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL collaboration -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: U01017; AAA81360.1; -.
DR EMBL: D50617; BAA09258.1; -.
DR PIR: S56274; S56274.
DR SGD: S0001915; FABI.
DR GO: GO:0000285; F:1-phosphatidylinositol-3-phosphate 5-kinase. . .; IDA.
DR GO: GO:0006644; P:phospholipid metabolism; IGI.
DR GO: GO:0007033; P:vacuole organization and biogenesis; IMP.
DR InterPro: IPR002498; PIP5K.
DR InterPro: IPR003305; Znf_FYVE.
DR Pfam: PF01363; FYVE; 1.
DR Pfam: PF01504; PIP5K; 1.
DR SMART: SM00064; FYVE; 1.
DR SMART: SM00330; PIP5K; 1.
DR PROSITE: PS50178; ZF_FYVE; 1.
KM Transferase; Kinase; Zinc-finger; Zinc.
FT ZN_FING 240 299 FYVE-TYPE.
FT DOMAIN 393 397 POLY-PRO.
FT DOMAIN 571 590 POLY-ASN.
FT DOMAIN 1808 1811 POLY-THR.
FT DOMAIN 1891 1897 POLY-GLN.
FT CONFLICT 2275 2275 R -> W (IN REF. 2).
SQ SEQUENCE 2278 AA; 257417 MW; 1A0A30E13165DE41 CRC64;

Query Match 1.2%; Score 15; DB 1; Length 2278;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 742 NNNNNNNNNNNNNN 756
Db 576 NNNNNNNNNNNNNN 590

RESULT 28
CTK2_XENLA STANDARD; PRT; 643 AA.
ID CTK2_XENLA
AC P79955;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cardoxy-terminal kinesin 2 (XCTK2).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_Taxid=8355;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=OVARY;
RX MEDLINE=97201465; PubMed=9049251;
RA Walczak C.E., Verma S., Mitchison T.J.;
RT "XCTK2: a kinesin-related protein that promotes mitotic spindle
RT assembly in Xenopus laevis egg extracts.";
RL J. Cell Biol. 136:859-870(1997).
CC -1- FUNCTION: PROMOTES MITOTIC SPINDLE ASSEMBLY.
CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS: A SMALL GLOBULAR
CC N-TERMINAL, A CENTRAL ALPHA-HELICAL COILED COIL AND A LARGE
CC GLOBULAR C-TERMINAL WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY
CC (IT HYDROLYZES ATP AND BINDS MICROTUBULES).
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD
CC SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: U82809; AAB40402.1; -.
CC HSSP: P17119; 3KAR.
DR InterPro: IPR001752; kinesin_motor.

```


DR Pfam; PF00225; kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00441; KINESIN MOTOR DOMAIN; 1.
 DR PROSITE; PS50067; KINESIN MOTOR DOMAIN; 1.
 KW Motor protein; Microtubules; ATP-binding; Coiled coil.
 FT DOMAIN 1 116 GLOBULAR.
 FT DOMAIN 117 296 COILED COIL (POTENTIAL).
 FT DOMAIN 291 564 KINESIN-MOTOR.
 FT NP BIND 386 393 ATP (POTENTIAL).
 SQ SEQUENCE 643 AA; 71948 MW; D3D92D08B8E5057 CRC64;

Query Match 1.3%; Score 14; DB 1; Length 643;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 FAYGQTSGKTYM 105
 DB 383 FAYGQTSGKTYM 396

RESULT 29
 NCD DROME STANDARD; PRT; 700 AA.
 ID NCD DROME
 AC P20480; Q9VAG8;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Claret segregational protein.
 GN NCD OR CA(ND) OR CG7831.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCB1_TaxID=7227;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Cealiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananthakrishnan P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan R.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-D., Andrews-Pfannkoch C., Baldwin D.,
 RA Baillie R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokov D., Borchen M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cusley S., Dahlke C., Davenport L.B., Davies P.,
 RA DePamphilis M., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Houston D., Houston K.A., Howard T.J., Wei M.-H., Ibegwam C.,
 RA Jalali N.D., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleab J.M.,

RA Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Styrkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN SEQUENCE FROM N.A.
 RC STRAIN=Berkley; TISSUE=Embryo;
 RP Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Fattan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Pacleab J., Paragas V., Park S., Phuenanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Cealiker S.;
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 RN SEQUENCE OF 16-700 FROM N.A.
 RP MEDLINE=90275618; PubMed=2140958;
 RX McDonald H.B., Goldstein L.S.B.;
 RT Identification and characterization of a gene encoding a
 RT kinesis-like protein in Drosophila.";
 RL Cell 61:991-1000(1990).
 RN MOTOR DIRECTIONALITY.
 RP MEDLINE=91043032; PubMed=2146510;
 RX Walker R.A., Salmon E.D., Endow S.A.;
 RT "The Drosophila claret segregation protein is a minus-end directed
 RT motor molecule.";
 RL Nature 347:780-782(1990).
 RN CHARACTERIZATION.
 RP MEDLINE=9415838; PubMed=8112290;
 RX Lockhart A., Cross R.A.;
 RT "Origins of reversed directionality in the ncd molecular motor.";
 RL EMBO J. 13:751-757(1994).
 RN MUTANT ALLELE NCD(D).
 RP MEDLINE=91122049; PubMed=1825056;
 RX Komar D.J., Horne A.S., Endow S.A.;
 RT "Separation of meiotic and mitotic effects of claret
 RT non-disjunctional on chromosome segregation in Drosophila.";
 RL EMBO J. 10:419-424(1991).
 RN CHARACTERIZATION OF MUTANT ALLELE NCD(D).
 RP MEDLINE=96283629; PubMed=8670831;
 RX Moore J.D., Song H., Endow S.A.;
 RT "A point mutation in the microtubule binding region of the Ncd motor
 RT protein reduces motor velocity.";
 RL EMBO J. 15:3306-3314(1996).
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 335-700.
 RP MEDLINE=96195067; PubMed=8606780;
 RX Sablin E.P., Kull F.J., Cooke R., Vale R.D., Fletcher R.J.;
 RT "Crystal structure of the motor domain of the kinesin-related motor
 RT ncd.";
 RL Nature 380:555-559(1996).
 CC -1- FUNCTION: NCD IS REQUIRED FOR NORMAL CHROMOSOMAL SEGREGATION IN
 CC MEIOSIS, IN FEMALES, AND IN EARLY MITOTIC DIVISIONS OF THE EMBRYO.
 CC THE NCD MOTOR ACTIVITY IS DIRECTED TOWARD THE MICROTUBULE'S MINUS
 CC END.
 CC -1- MISCELLANEOUS: NCD(D) IS A MUTANT ALLELE THAT SHOWS ABNORMAL
 CC CHROMOSOMAL SEGREGATION.
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD
 CC SUBFAMILY.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X52814; CAA36998.1; -
DR EMBL; M33932; AAA28716.1; -
DR EMBL; AE003771; AAF56942.1; -
DR EMBL; AY058596; AAL13825.1; -
DR EMBL; X57475; CAA40713.1; -
DR PIR; S09748; S09748.
DR PDB; 1CZ7; 05-NOV-99.
DR PDB; 2NCD; 09-JUN-99.
DR FlyBase; FBgn0002924; ncd.
DR GO; GO:0007100; P:centrosome separation; IGI.
DR GO; GO:0000212; P:meiotic spindle assembly; NAS.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin_1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
KW Meiosis; Mitosis; 3D-structure.
FT DOMAIN 196 346 COILED COIL (POTENTIAL).
FT NP BIND 347 700 KINESIN-MOTOR (BY SIMILARITY).
FT NP BIND 434 441 ATP (BY SIMILARITY).
FT MUTAGEN 556 556 V->F; IN NCD(D); REDUCES MOTOR VELOCITY.
FT CONFLICT 697 697 S->N (IN REF. 1).
SQ SEQUENCE 700 AA; 77473 MW; ADE043CBCE7FD561 CRC64;

Query Match 1.1%; Score 14; DB 1; Length 700;
Best Local Similarity 100.0%; Pred. No. 3.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 92 FAYGOTSGKTYTM 105
DB 431 FAYGOTSGKTYTM 444

RESULT 30
ID_KCM1_XENLA STANDARD; PRT; 730 AA.
AC 091636;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Kinesin central motor 1 (XKCM1).
GN KCM1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=96140638; PubMed=8548824;
RA Walczak C.E., Mitchison T.J., Desai A.;
RT "XKCM1: a Xenopus kinesin-related protein that regulates microtubule
RT dynamics during mitotic spindle assembly.";
RL Cell 84:37-47(1996).
RN [2]
RP REVISION TO 516.
RC TISSUE=Ovary;
RA Walczak C.E.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: REGULATES MICROTUBULE DYNAMICS DURING MITOTIC SPINDLE
CC ASSEMBLY. REQUIRED FOR BOTH ESTABLISHMENT AND MAINTENANCE OF
CC MITOTIC SPINDLES. MAY ACT DIRECTLY ON MICROTUBULES, CAUSING A
CC DESTABILIZATION AND EVENTUAL DEPOLYMERIZATION OF THE MICROTUBULE.
CC -I- SUBUNIT: Homodimer (Potential).

CC -I- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC -I- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. MCAK/KIF2
CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U36485; AAC59743.2; -
DR HSSP; P17119; 3KAR.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin_1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
KW Domain 1 256 GLOBULAR (POTENTIAL).
FT DOMAIN 257 598 KINESIN-MOTOR (BY SIMILARITY).
FT NP BIND 599 730 COILED COIL (POTENTIAL).
FT NP BIND 352 359 ATP (POTENTIAL).
SQ SEQUENCE 730 AA; 82585 MW; 25C31C187E491523 CRC64;

Query Match 1.1%; Score 14; DB 1; Length 730;
Best Local Similarity 100.0%; Pred. No. 3.9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 263 INRSLLAKECIRA 276
DB 520 INRSLLAKECIRA 533

RESULT 31
ID_ATK2_ARATH STANDARD; PRT; 745 AA.
AC P46864; O9T047;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kinesin 2 (kinesin-like protein B).
GN ATK2 OR KATB OR ATG427180 OR T24A18.130.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=94355659; PubMed=8075402;
RA Mitsui H., Nakatani K., Yamaguchi-Shinozaki K., Shinozaki K.,
RA Nishikawa K., Takahashi H.;
RT "Sequencing and characterization of the kinesin-related genes katB
RT and katC of Arabidopsis thaliana.";
RL Plant Mol. Biol. 25:865-876 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueler C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansoerge W., Brandt P., Grivell U.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,
RA Kreis M., Deleseny M., Puigdomenech P., Watson M., Schmittheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hohenseil U., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Roben J., Vandenbusche F.,
RA Van der Schueren J., Gymnopoulos B., Chiang Y.-J., Vandenbusche F.,
RA Breken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,

RA Weitzenecker T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA Holzner E., Brandt A., Peters S., van Staveren M., Dirks W.,
RA Moilman P., Klein Lankhorst R., Rose M., Haut J., Koetser P.,
RA Bernier S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLeay K., Mayes R.,
RA Pettit A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehert T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Wellner Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Granderath K., Danner D., Herzi A.,
RA Neumann S., Argirion A., Vitale D., Liguori R., Pirvandi E.,
RA Massenet O., Quigley F., Clabaud G., Mündlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chefor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Torres A.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
RA Frisman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Hevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedria N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stenking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Lattelle P., Courteney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Giesel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Jochu C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong Y., Preston R., Vil D., Shekhar M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
RA Granat S., Shoddy N., Hasegawa A., Hameed A., Lohi M., Johnson A.,
RA Chen E., Maria M., Martienssen R., McCombie W.R.,
RT "Sequence and analysis of chromosome 4 of the plant *Arabidopsis*
RT *thaliana*." Nature 402:769-777(1999).
CC -1- FUNCTION: POSSIBLE ROLE IN MITOSIS.
CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS; A SMALL GLOBULAR
CC N-TERMINAL, A CENTRAL ALPHA-HELICAL COILED COIL AND A LARGE
CC GLOBULAR C-TERMINAL WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY
CC (IT HYDROLYZES ATP AND BINDS MICROTUBULES).
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD
CC SUBFAMILY.
CC -1- CAUTION: Ref.2 sequences differ from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D21137, BAA04673.1; -
CC EMBL: AL035680, CAB38848.1; ALT_SEQ.
CC EMBL: AL015566, CAB79573.1; ALT_SEQ.
CC PIR: T06048; T06048.
CC HSSP: P17119, 3KAR.
CC InterPro: IPR001752, kinesin_motor.
CC DR Pfam: PF00225, kinesin_1.
CC DR PRINTS: PRO0380, KINESINHEAVY.
CC DR SMART: SM00129, KISC.1.
CC DR PROSITE: PS00411, KINESIN MOTOR DOMAIN; 1.
CC DR PROSITE: PS50067, KINESIN MOTOR DOMAIN; 1.
CC KW Motor protein; Microtubules; ATP-binding; Coiled coil;
CC Multigene family.
CC FT DOMAIN 1 35 GLOBULAR.
CC FT DOMAIN 36 384 COILED COIL.
CC FT DOMAIN 385 718 KINESIN-MOTOR (BY SIMILARITY).
CC FT NP_BIND 472 479 ATP (POTENTIAL).
CC SQ SEQUENCE 745 AA; 84359 MW; 6DB352FDF4FD7BC1 CRC64;

Query Match 1.1%; Score 14; DB 1; Length 745;
Best Local Similarity 100.0%; Pred. No. 4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 469 FAYGONGSGKTYTM 482
92 FAYGONGSGKTYTM 105
|||||
ATK3 ARATH STANDARD; PRT; 754 AA.
ID ATK3 ARATH
AC P46875; Q9FH38;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE kinesin 3 (kinesin-like protein C).
GN ATK3 OR KATC OR ATSG54670 OR K5F14.1 OR MEB17.18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosidie II; Brassicales; Brassicaceae; Arabidopsie.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=94355659; PubMed=8075402;
RA Mitsui H., Nakatani K., Yamaguchi-Shinozaki K., Shinozaki K.,
RT Nishikawa K., Takahashi H.,
RT "Sequencing and characterization of the kinesin-related genes *katb*
RT and *kac* of *Arabidopsis thaliana*." Plant Mol. Biol. 25:865-876(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RT Tabata S.,
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones." J. Biol. Res. 7:31-63(2000).
RL DNA Res. 7:31-63(2000).
CC -1- FUNCTION: POSSIBLE ROLE IN MITOSIS.
CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS; A SMALL GLOBULAR
CC N-TERMINAL, A CENTRAL ALPHA-HELICAL COILED COIL AND A LARGE
CC GLOBULAR C-TERMINAL WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY
CC (IT HYDROLYZES ATP AND BINDS MICROTUBULES).
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD
CC SUBFAMILY.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D21138, BAA04674.1; -
CC EMBL: AB022214, BAB09933.1; ALT_SEQ.
CC PIR: S48020; S48020.
CC HSSP: P17119, 3KAR.
CC InterPro: IPR001752, kinesin_motor.
CC DR Pfam: PF00225, kinesin_1.
CC DR PRINTS: PRO0380, KINESINHEAVY.
CC DR SMART: SM00129, KISC.1.
CC DR PROSITE: PS00411, KINESIN MOTOR DOMAIN; 1.
CC DR PROSITE: PS50067, KINESIN MOTOR DOMAIN; 1.
CC KW Motor protein; Microtubules; ATP-binding; Coiled coil; Mitosis;
CC Multigene family.
CC FT DOMAIN 1 45 GLOBULAR.


```

FT DOMAIN 46 393 COILED COIL.
FT DOMAIN 398 727 KINESIN-MOTOR (BY SIMILARITY).
FT NP BIND 481 488 ATP (POTENTIAL).
SQ SEQUENCE 754 AA; 85030 MW; 76091CD585D9C531 CRC64;

Query Match 1.1%; Score 14; DB 1; Length 754;
Best Local Similarity 100.0%; Pred. No. 4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 92 PAYGOTSGKTYM 105
DB 478 PAYGOTSGKTYM 491

RESULT 33
ATK1_ARATH
ID ATK1_ARATH STANDARD; PRT; 793 AA.
AC Q07970;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kinesin 1 (Kinesin-like protein A).
GN ATK1 OR KATA OR AT4G21270 OR F7J7.210.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI TaxID=3702;
RX MEDLINE=93261419; PubMed=8492804;
RA Mitsui H., Yamaguchi-Shinozaki K., Shinozaki K., Nishikawa K.,
RA Takahashi H.;
RT "Identification of a gene family (kat) encoding kinesin-like proteins
RT in Arabidopsis thaliana and the characterization of secondary
RT structure of KAT1."
RL Mol. Genet. 238:362-368(1993).

[2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schaefer C., Stiekema W., Entian K.-D., Terryn N.,
RA Pohl T., Duescherhoft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
RA Weichenberger M., de Simone V., Obermayer B., Maehne R., Mueller M.,
RA Kreis M., Delsen W., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hehseisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenhussche F.,
RA Bieken M., Welfjens I., Voet M., Baetiaens I., Aert R., Deleor E.,
RA Holzernegger T., Bothe G., Rampsperger U., Hilbert H., Braun M.,
RA Weitzer E., Brandt A., Peters S., van Staveren M., Dirse W.,
RA Moeljan P., Klein lankhorst R., Rose M., Hauf J., Koetter P.,
RA Benneiser S., Hempel S., Feldausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Leonard N., Mcay K., Mayes R.,
RA Petrecca A., Rajeandrem M.A., Lyne M., Benes V., Reckmann S.,
RA Borikova D., Bloeker H., Scharf M., Grimm M., Loehner T.-H.,
RA Doe S., de Haan M., Maaree A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Granderth K., Danner D., Herzi A.,
RA Neumann S., Agitlou A., Vitale D., Lignori R., Pivavandi E.,
RA Masenat O., Quigley F., Clabaud G., Muendlein A., Feilber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chedors T., Weber N., Vandenbol M., Bagnies M., Terol J., Torres A.,
RA Perez-Perez A., Furnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijmen D., Haase D., Lemcke S., Scholier P., Heber S., Francis C.,
RA Zaccaria P., Bevan W., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,

```

```

RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Marks E., Dante W., Pavin K., Hillier L.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Gelsel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Dione K., Cotton M., Joshi C.,
RA Antonov B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekhar M., Maturo A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Haegawa A., Hamed A., Lohi M., Johnson A.,
RA Chen E., Marra M., Martienssen R., McCombe W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana."
RL Nature 402:769-777(1999).

[3]
RP FUNCTION.
RX PubMed=12112142;
RA Marcus A.I., Ambrose J.C., Blackley L., Hancock W.O., Cyr R.J.;
RT "Arabidopsis thaliana protein, ATK1, is a minus-end directed kinesin
RT that exhibits non-processive movement."
RL Cell Motil. Cytoskeleton 52:144-150(2002).

[4]
RN RN
RP FUNCTION.
RX PubMed=11973272;
RA Chen C., Marcus A., Li W., Hu Y., Calzada J.-P., Grossniklaus U.,
RA Cyr R.J., Ma H.;
RT "The Arabidopsis ATK1 gene is required for spindle morphogenesis in
RT male meiosis."
RL Development 129:2401-2409(2002).

-1- DOMAIN: Kinesin that supports microtubule movement in an ATP-
dependent manner and has a minus-end directed polarity. Plays a
crucial role in spindle morphogenesis in male Arabidopsis meiosis.
In ATK1 mutants, male meiosis is defective, producing an abnormal
number of microspores of variable sizes.

-1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS; A SMALL GLOBULAR
N-TERMINAL, A CENTRAL ALPHA-HELICAL COILED COIL, AND A LARGE
GLOBULAR C-TERMINAL WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY
(1T HYDROLATYSE ATP AND BINDS MICROTUBULES).

-1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD
SUBFAMILY.

-----
CC EMBL, D11371; BAA01972.1; -.
CC EMBL, AL021960; CAA17546.1; -.
CC EMBL, AL161554; CAB79127.1; -.
CC PIR, S34830; S34830.
CC HSSP, P17119; 3KAR.
CC InterPro: IPR001752; kinesin_motor.
CC Pfam, PF00225; kinesin_1.
CC SMART, PR00380; KINESINHEAVY.
CC PRINTS, SM00129; KISC; 1.
CC PROSITE, PS00411; KINESIN MOTOR DOMAIN1, 1.
CC PROSITE, PS00674; KINESIN MOTOR DOMAIN2, 1.
CC Motor protein; Microtubules; ATP-binding; Coiled coil; Meiosis;
CC Multigene family.
CC GLOBULAR.
CC COILED COIL.
CC KINESIN-MOTOR (BY SIMILARITY).
CC ATP (POTENTIAL).
SQ SEQUENCE 793 AA; 89047 MW; 85A195206D825C4E CRC64;

Query Match 1.1%; Score 14; DB 1; Length 793;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

QY      92 FAYGQTSGSKTYTM 105
DB      513 FAYGQTSGSKTYTM 526

RESULT 34
ID      KLP1_SCHPO          STANDARD;          PRT;          832 AA.
AC      092376; 042669;
DT      01-NOV-1997 (Rel. 35, Created)
DT      15-DEC-1998 (Rel. 37, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Kinesin-like protein 1.
GN      PGL1 OR SPAC3A11.14C.
OS      Schizosaccharomyces pombe (Fission Yeast).
OC      Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC      Schizosaccharomycetales; Schizosaccharomycetaceae;
OC      Schizosaccharomyces.
OX      NCBI_TaxID=4896;
ON      (1)
RP      SEQUENCE FROM N.A.
RC      STRAIN=972;
RX      MEDLINE=97053988; PubMed=8898367;
RT      Pidoux A.L., Ledizet M., Candé W.Z.;
RT      "Fission yeast pkl1 is a kinesin-related protein involved in mitotic
RT      spindle function."
RL      Mol. Biol. Cell 7:1639-1655 (1996).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=972;
RX      MEDLINE=21848401; PubMed=11859360;
RX      Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA      Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA      Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA      Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
RA      Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA      Holtz S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA      James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA      Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA      Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA      Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA      Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA      Taylor R.G., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA      Woodward J., Voiclaert G., Aert R., Robben J., Gymnopoulos B.,
RA      Wellens I., Vansireels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA      Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA      Borzym K., Langer I., Beck A., Lehach H., Reinhardt R., Pohl T.M.,
RA      Eger P., Zimmermann W., Wedler H., Wambutt R., Burnelle B.,
RA      Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA      Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA      Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA      Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA      Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA      Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA      Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT      "The genome sequence of Schizosaccharomyces pombe."
RL      Nature 415:871-880 (2002).
CC      -1- FUNCTION: MICROTUBULE-DEPENDENT MOTOR THAT IS INVOLVED IN
CC      MICROTUBULE ORGANIZATION IN THE MITOTIC SPINDLE.
CC      -1- SUBCELLULAR LOCATION: Nuclear.
CC      -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD
CC      SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; U63916; AAB88235.1; -
CC      DR      EMBL; 299260; CAB16389.1; -

```

```

DR      EMBL; 299260; CAB16389.1; -
DR      PIR; T38749; T38749.
DR      HSSP; P17119; 3KAR.
DR      GeneDB SPombe; SPAC3A11.14C; -.
DR      InterPro; IPR001752; kinesin_motor.
DR      Pfam; PF00225; kinesin; 1.
DR      SMART; SM00380; KINSEINHEAVY.
DR      SMART; SM00411; KINESIN MOTOR DOMAIN1; 1.
DR      PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
KM      Motor protein; Microtubules; ATP-binding; coiled coil;
KM      Nuclear protein.
FT      DOMAIN 193 266 COILED COIL (POTENTIAL).
FT      DOMAIN 284 305 COILED COIL (POTENTIAL).
FT      DOMAIN 344 481 COILED COIL (POTENTIAL).
FT      DOMAIN 482 832 KINESIN-MOTOR (BY SIMILARITY).
FT      NP_BIND 575 582 ATP (POTENTIAL).
FT      CONFLICT 343 343 E -> D (IN REF. 1).
SQ      SEQUENCE 832 AA; 96336 MW; 870CF7424B9F3E4 CRC64;

Query Match      1.1%; Score 14; DB 1; Length 832;
Best local Similarity 100.0%; Pred. No. 4.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      92 FAYGQTSGSKTYTM 105
DB      572 FAYGQTSGSKTYTM 585

RESULT 35
ID      CYAG_DICDI          STANDARD;          PRT;          858 AA.
AC      003101;
DT      01-JUL-1993 (Rel. 26, Created)
DT      01-JUL-1993 (Rel. 26, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Adenylate cyclase, germination specific (EC 4.6.1.1) (ATP
DE      pyrophosphate-lyase) (Adenylyl cyclase).
GN      ACQA OR ACG.
OS      Dictyostelium discoideum (Slime mold).
OC      Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX      NCBI_TaxID=44689;
ON      (1)
RP      SEQUENCE FROM N.A.
RC      MEDLINE=92233467; PubMed=1348970;
RX      Pitt G.S., Milona N., Bortels J., Lin K.C., Reed R.R.,
RA      Devreotes P.N.;
RA      "Structurally distinct and stage-specific adenylyl cyclase genes play
RA      different roles in Dictyostelium development."
RL      Cell 69:305-315 (1992).
RN      [2]
RP      3D-STRUCTURE MODELING OF 387-543.
RX      MEDLINE=98054247; PubMed=9391039;
RA      Liu Y., Ruoho A.E., Rao V.D., Hurley J.H.;
RT      "Catalytic mechanism of the adenylyl and guanylyl cyclases: modeling
RT      and mutational analysis."
RL      Proc. Natl. Acad. Sci. U.S.A. 94:13414-13419 (1997).
CC      -1- FUNCTION: ACG HAS A LARGE EXTRACELLULAR DOMAIN WHICH MAY BE
CC      INVOLVED IN THE RECOGNITION OF AN EXTRACELLULAR SIGNAL PRESENT
CC      DURING GERMINATION, LEADING TO ACTIVATION OR INHIBITION OF CAMP
CC      SYNTHESIS BY THE CYTOPLASMIC DOMAIN.
CC      -1- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC      -1- ENZYME REGULATION: INSENSITIVE TO GUANINE NUCLEOTIDES.
CC      -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC      -1- DEVELOPMENTAL STAGE: AFTER FRUITING BODIES HAVE BEEN FORMED AND
CC      DURING GERMINATION.
CC      -1- SIMILARITY: Belongs to the adenylyl cyclase class-4/guanylyl
CC      cyclase family.
CC      -1- SIMILARITY: Contains 1 CHASE domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its

```


use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M87278; AAA33164.1; --
DR PDB; 1AMW; 28-JAN-98.
DR DictyDb; DDo4007; acga.
DR InterPro; IPR006189; CHASE.
DR InterPro; IPR01054; G_cyclase.
DR Pfam; PF03924; CHASE; 1.
DR SMART; SM00044; CYCC; 1.
DR PROSITE; PS05039; CHASE; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE; PS0125; GUANYLATE_CYCLASES_2; 1.
KW Lyase; CAMP biosynthesis; Transmembrane; Germination; Signal-anchor; 3D-structure.
FT DOMAIN 1 18 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 19 41 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT DOMAIN 42 858 (POTENTIAL).
FT DOMAIN 86 317 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 396 526 CHASE.
FT DOMAIN 738 848 GUANYLATE CYCLASE.
FT DOMAIN 835 848 ASN-RICH.
FT STRAND 391 398 POLY-ASN.
FT STRAND 403 408
FT HELIX 415 432
FT TURN 433 434
FT STRAND 439 443
FT TURN 444 445
FT STRAND 446 451
FT HELIX 462 473
FT TURN 474 474
FT TURN 479 482
FT STRAND 492 492
FT STRAND 495 500
FT STRAND 504 504
FT STRAND 511 511
FT STRAND 517 527
FT HELIX 534 534
FT STRAND 538 542
FT HELIX 542 542
SQ SEQUENCE 858 AA; 98413 MW; B483FB255289E65D CRC64;

Query Match 1.1%; Score 14; DB 1; Length 858;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 742 NNNNNNNNNNNNNN 755
DB 835 NNNNNNNNNNNNNN 848

RESULT 36
KINL_LEICH STANDARD; PRT; 955 AA.
ID KINL_LEICH
AC P46865;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Kinesin-like protein K39 (Fragment).
GN KIN.
OS Leishmania chagasi.
OC Eukaryota; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=44271;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MHOM/BR/82 / Isolate BA-2;
RX MEDLINE=93133867; PubMed=8421715;
RA Burns J.M. Jr., Shreffler W.G., Benson D.R., Ghallib H.W., Badaro R., Reed S.G.;
RT "Molecular characterization of a kinesin-related antigen of

Leishmania chagasi that detects specific antibody in African and American visceral leishmaniasis.";
Proc Natl. Acad. Sci. U.S.A. 90:775-779(1993).
CC -1- DEVELOPMENTAL STAGE: PREDOMINANT IN AMASTIGOTES.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; L07879; AAA29254.1; --
DR PIR; A47334; A47334.
DR HSSP; P17119; 3KAR.
DR InterPro; IPR01752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PRO0380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil; Repeat.
FT DOMAIN 1 399 KINESIN-MOTOR (BY SIMILARITY).
FT NP_BIND 426 955
FT NP_BIND 122 129 COILED COIL (POTENTIAL).
FT DOMAIN 704 742 ATP (POTENTIAL).
FT REPEAT 704 742 7 X 39 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 743 781 1.
FT REPEAT 782 820 2.
FT REPEAT 821 859 3.
FT REPEAT 860 898 4.
FT REPEAT 899 937 5.
FT REPEAT 938 955 6.
FT REPEAT 955 955 7 (PARTIAL).
SQ SEQUENCE 955 AA; 106168 MW; 8CA76815B84C69 CRC64;

Query Match 1.1%; Score 14; DB 1; Length 955;
Best Local Similarity 100.0%; Pred. No. 5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 92 FAYGQTSGKTYTM 105
DB 119 FAYGQTSGKTYTM 132

RESULT 37
R115_YEAST STANDARD; PRT; 1770 AA.
ID R115_YEAST
AC P43565;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase RIM15 (EC 2.7.1.-).
GN RIM15 OR TAK1 OR YPL033C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M., Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K., Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from Saccharomyces cerevisiae";
RL Nat. Genet. 10:261-268(1995).
RN [2]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=97265402; PubMed=9111339;

RA vidan S., Mitchell A.P.;
 RT "stimulation of yeast meiotic gene expression by the
 RT glucose-repressible protein kinase Rim15p."; Mol.
 RN Cell. Biol. 17:2688-2697(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Reinhardt A., Buerckert N., Bolter T., Wiemken A., de Virgilio C.;
 RT "Saccharomyces cerevisiae CAMP-dependent protein kinase controls entry
 RT into stationary phase through the Rim15p protein kinase."; Mol.
 RN Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN THE ACTIVATION OF A MEIOTIC GENES ACTIVATION
 CC PATHWAY.
 CC -1- PTM: AUTOPHOSPHORYLATED.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC STRONG, TO S.POMBE CEXL.
 CC -1- SIMILARITY: Contains 1 response regulatory domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D50617; BAA09206.1; -;
 DR EMBL; U83459; AAB64088.1; -;
 DR EMBL; AJ001030; CAA04486.1; -;
 DR PIR; S56221; S56221.
 DR SGD; S0001861; KIM15.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
 DR GO; GO:0040020; P:regulation of meiosis; IMP.
 DR GO; GO:0006950; P:response to stress; IMP.
 DR InterPro; IPR000961; Kinase.C.
 DR InterPro; IPR000719; Prot_Kinase.
 DR InterPro; IPR001789; Response_reg.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00069; kinase; 1.
 DR Pfam; PF00072; response_reg; 1.
 DR ProDom; PD000001; Prot_Kinase; 2.
 DR ProDom; PD000039; Response_reg; 1.
 DR SMART; SM00448; REC; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
 DR PROSITE; PS00111; PROTEIN KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
 DR PROSITE; PS00110; RESPONSE REGULATORY; 1.
 DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding;
 KM Phosphorylation; Meiosis.
 FT DOMAIN 794 1254 PROTEIN KINASE.
 FT DOMAIN 1636 1750 RESPONSE REGULATORY.
 FT DOMAIN 343 358 POLY-ASN.
 FT NP_BIND 800 808 POLY-SER.
 FT BINDING 823 823 ATP (BY SIMILARITY).
 FT ACT_SITE 918 918 ATP (BY SIMILARITY).
 FT DOMAIN 975 980 BY SIMILARITY.
 FT DOMAIN 1213 1218 POLY-ASN.
 FT DOMAIN 1386 1391 POLY-GLU.
 FT DOMAIN 1391 POLY-THR.
 SQ SEQUENCE 1770 AA; 196530 MW; DC1064825000FAFP CRC64;

Query Match 1.1%; Score 14; DB 1; Length 1770;
 Best Local Similarity 100.0%; Pred. No. 8.9e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 742 NNNNNNNNNNNNNNNN 755
 |||||
 Db 345 NNNNNNNNNNNNNNNN 358

RESULT 38
 YMBD YEAST

ID YMBD YEAST STANDARD; PRT; 201 AA.
 AC P38429; Q03501;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hypothetical 23.0 kDa protein in TPS3-IP2 intergenic region (ORF2).
 GN YMR263W OR YMR156.05.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Manning A.M., Rosenblum C.L., Beaudet A.L.;
 RT Submitted (Jun-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX PubMed=9169872;
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
 RA Jags K., Iye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
 RA Rice P., Skelton J., Walsh S., Whitehead S., Barrall B.G.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
 RT XIII.";
 RL Nature 387:90-93(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M68172; AAA5222.1; -;
 DR EMBL; Z49260; CAA89246.1; -;
 DR PIR; S54475; S54475.
 DR SGD; S0004876; SAP30.
 DR GO; GO:0000118; C:histone deacetylase complex; IPI.
 DR GO; GO:0004407; F:histone deacetylase activity; IPI.
 DR GO; GO:0006347; P:chromatin silencing at HML and HMR (sensu S. .); IMP.
 DR GO; GO:0000183; P:chromatin silencing at telomeric DNA (rdNA); IMP.
 DR GO; GO:0006348; P:chromatin silencing at telomere; IMP.
 KM Hypothetical protein.
 FT DOMAIN 26 58 ASN-RICH.
 FT CONFLICT 31 31 C -> S (IN REF. 1).
 FT CONFLICT 48 48 N -> S (IN REF. 1).
 SQ SEQUENCE 201 AA; 23026 MW; 6C0D2734C50DD72E CRC64;

Query Match 1.0%; Score 13; DB 1; Length 201;
 Best Local Similarity 100.0%; Pred. No. 0.00011;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 502 NNNNNNNNNNNNNNS 514
 |||||
 Db 39 NNNNNNNNNNNNNNS 51

RESULT 39
 HKLB LYCES
 ID HKLB LYCES STANDARD; PRT; 426 AA.
 AC Q22300;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein knotted-1 like LFT12.
 GN LFT12.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
 OC Asteridae; lamiales; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;


```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. VENT Cherry;
RX MEDLINE=98145476; PubMed=9484482;
RA Janssen B.J., Williams A., Chen J.J., Mathern J., Hake S., Sinha N.;
RT "Isolation and characterization of two knotted-like homeobox genes
  from tomato.";
RL Plant Mol. Biol. 36:417-425(1998).
CC -1- FUNCTION: MAY HAVE A ROLE TO PLAY IN FORMATIVE EVENTS IN OVULE AND
  EMBRYO MORPHOGENESIS.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED IN THE MATURE PLANT.
CC -1- SIMILARITY: BELONGS TO THE TALE/KNOX HOMEBOX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AF000142; AAC49918.1; -.
DR PIR, T04318; T04318.
DR TRANSEAC; T04055; -.
DR InterPro; IPR005539; ELK.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR005540; KNOX1.
DR InterPro; IPR005541; KNOX2.
DR Pfam; PF03789; ELK; 1.
DR Pfam; PF03790; KNOX1; 1.
DR Pfam; PF03791; KNOX2; 1.
DR Prodom; PD00010; Homeobox; 1.
DR SMART; SMO0389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
DR KRN-binding; Homeobox; Nuclear protein.
FT DOMAIN 15 24 POLY-GLN.
FT DOMAIN 69 76 POLY-ALA.
FT DOMAIN 140 152 POLY-ASN.
FT DOMAIN 283 287 POLY-ASP.
FT DOMAIN 325 348 ELK DOMAIN.
FT DNA BIND 349 411 HOMEBOX (TALE-TYPE).
SO SEQUENCE 426 AA; 47581 MW; 5B52B9E0A34A86C CRC64;

Query Match 1.0%; Score 13; DB 1; Length 426;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 742 NNNNNNNNNNNN 754
Db 140 NNNNNNNNNNNN 152

RESULT 40
KLP3 SCHPO STANDARD; PRT; 554 AA.
AC Q9US60; Q9US61;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Kinesin-like protein 3 (Kinesin-related protein 1).
GN KLP3 OR KRP1 OR SPAC183.07.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.
RC STRAIN=972;
RX MEDLINE=20109189; PubMed=10641037;
RA Brazier S.-C.W., Williams H.P., Chapell T.G., Cande W.Z.;

```

```

RT "A fission yeast kinesin affects Golgi membrane recycling.";
RL Yeast 16:149-166(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=22127603; PubMed=1212578;
RA Jeong J.W., Rhee D.K., Cho S.Y., Hae K.L., Kim D.U., Won M., Kim H.B.;
RT "Cloning and characterization of the kinesin-related protein, Krip1,
  in Schizosaccharomyces pombe.";
RL Mol. Cells 13:389-398(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream A., Lyne M., Lyne R., Stewart A.,
  Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
  Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
  Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
  Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
  Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
  James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
  Mooney P., Mounle S., Mungall K., Murphy L., Niblett D., Odell C.,
  Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
  Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
  Skelton J., Simmonds M., Squares R., Stevens K.,
  Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
  Woodward J., Voicikert G., Aert R., Robben J., Grymoprez B.,
  Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
  Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
  Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
  Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
  Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,
  Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
  Lucas M., Rocher M., Gallardin C., Tallada V.A., Garzon A., Rhode G.,
  Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
  Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,
  Riera L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
  Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- FUNCTION: CYTOPLASMIC MOTOR THAT COULD PLAY A ROLE IN GOLGI
  MEMBRANE RECYCLING.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AF154055; AAF14525.1; -.
DR EMBL, AF156966; AAF22609.1; -.
DR EMBL, AF247188; AAF81205.1; -.
DR EMBL, AL157734; CAB75775.1; -.
DR PIR, T50118; T50118.
DR HSSP; P33176; I862.
DR GeneDB; SPombe; SPAC183.07; -.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SMO0129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Coiled coil; Motor protein.
FT DOMAIN 1 332 KINESIN-MOTOR (BY
  SIMILARITY).
FT NP_BIND 446 473 COILED COIL (POTENTIAL).
FT NP_BIND 84 91 ATP (POTENTIAL).
FT NP_BIND 233 240 ATP (POTENTIAL).
FT CONFLICT 57 57 O -> L (IN REF. 1).

```


SO SEQUENCE 554 AA; 61938 MW; B4533BD095C49CHD CRC64;

Query Match 1.0%; Score 13; DB 1; Length 554;

Best Local Similarity 100.0%; Pred. No. 0.00028;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 93 AYGQTSGSGKTYTM 105

Db 82 AYGQTSGSGKTYTM 94

Search completed: October 2, 2003, 16:32:52
Job time : 51 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: October 2, 2003, 16:21:36 ; Search time 94 Seconds
(without alignments)
2174.892 Million cell updates/sec

Title: US-10-006-780-2

Perfect score: 1288

Sequence: 1 MNSKIKVVRKRLSELEKK.....KKLVODNINSMNDNNHKK 1288

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0, ...

Searched: 1107863 seqs, 158726573 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database:

A.Geneseq_19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	1.3	190	14	Staphylococcus ent
2	17	1.3	190	14	Staphylococcus ent
3	17	1.3	190	14	Staphylococcus ent
4	17	1.3	190	14	Staphylococcus ent
5	17	1.3	190	14	Staphylococcus ent
6	17	1.3	190	14	Staphylococcus ent
7	17	1.3	190	14	Staphylococcus ent
8	17	1.3	190	14	Staphylococcus ent
9	17	1.3	190	14	Staphylococcus ent

10	17	1.3	190	14	Staphylococcus ent
11	17	1.3	190	14	Staphylococcus ent
12	17	1.3	190	14	Staphylococcus ent
13	17	1.3	190	14	Staphylococcus ent
14	17	1.3	190	14	Staphylococcus ent
15	17	1.3	190	14	Staphylococcus ent
16	17	1.3	190	14	Staphylococcus ent
17	17	1.3	190	14	Staphylococcus ent
18	17	1.3	190	14	Staphylococcus ent
19	17	1.3	190	14	Staphylococcus ent
20	17	1.3	190	14	Staphylococcus ent
21	17	1.3	190	14	Staphylococcus ent
22	17	1.3	190	14	Staphylococcus ent
23	17	1.3	190	14	Staphylococcus ent
24	17	1.3	190	14	Staphylococcus ent
25	17	1.3	190	14	Staphylococcus ent
26	17	1.3	190	14	Staphylococcus ent
27	17	1.3	190	14	Staphylococcus ent
28	17	1.3	190	14	Staphylococcus ent
29	17	1.3	190	14	Staphylococcus ent
30	17	1.3	190	14	Staphylococcus ent
31	17	1.3	190	14	Staphylococcus ent
32	17	1.3	190	14	Staphylococcus ent
33	17	1.3	190	14	Staphylococcus ent
34	17	1.3	190	14	Staphylococcus ent
35	17	1.3	190	14	Staphylococcus ent
36	17	1.3	190	14	Staphylococcus ent
37	17	1.3	190	14	Staphylococcus ent
38	17	1.3	190	14	Staphylococcus ent
39	17	1.3	190	14	Staphylococcus ent
40	17	1.3	190	14	Staphylococcus ent
41	17	1.3	190	14	Staphylococcus ent
42	17	1.3	190	14	Staphylococcus ent
43	17	1.3	190	14	Staphylococcus ent
44	17	1.3	190	14	Staphylococcus ent
45	17	1.3	190	14	Staphylococcus ent
46	17	1.3	190	14	Staphylococcus ent
47	17	1.3	190	14	Staphylococcus ent
48	17	1.3	190	14	Staphylococcus ent
49	17	1.3	190	14	Staphylococcus ent
50	17	1.3	190	14	Staphylococcus ent
51	17	1.3	190	14	Staphylococcus ent
52	17	1.3	190	14	Staphylococcus ent
53	17	1.3	190	14	Staphylococcus ent
54	17	1.3	190	14	Staphylococcus ent
55	17	1.3	190	14	Staphylococcus ent
56	17	1.3	190	14	Staphylococcus ent
57	17	1.3	190	14	Staphylococcus ent
58	17	1.3	190	14	Staphylococcus ent
59	17	1.3	190	14	Staphylococcus ent
60	17	1.3	190	14	Staphylococcus ent
61	17	1.3	190	14	Staphylococcus ent
62	17	1.3	190	14	Staphylococcus ent
63	17	1.3	190	14	Staphylococcus ent
64	17	1.3	190	14	Staphylococcus ent
65	17	1.3	190	14	Staphylococcus ent
66	17	1.3	190	14	Staphylococcus ent
67	17	1.3	190	14	Staphylococcus ent
68	17	1.3	190	14	Staphylococcus ent
69	17	1.3	190	14	Staphylococcus ent
70	17	1.3	190	14	Staphylococcus ent
71	17	1.3	190	14	Staphylococcus ent
72	17	1.3	190	14	Staphylococcus ent
73	17	1.3	190	14	Staphylococcus ent
74	17	1.3	190	14	Staphylococcus ent
75	17	1.3	190	14	Staphylococcus ent
76	17	1.3	190	14	Staphylococcus ent
77	17	1.3	190	14	Staphylococcus ent
78	17	1.3	190	14	Staphylococcus ent
79	17	1.3	190	14	Staphylococcus ent
80	17	1.3	190	14	Staphylococcus ent
81	17	1.3	190	14	Staphylococcus ent
82	17	1.3	190	14	Staphylococcus ent

83	14	1.1	375	21	AAAG31717	Arabidopsis thalia	156	12	0.9	729	22	ABBS9500	Drosophila melanog
84	14	1.1	375	21	AAAY70952	Arabidopsis SERK 1	157	12	0.9	761	22	ABBR1561	Drosophila melanog
85	14	1.1	458	21	ABP73738	Candida albicans e	158	12	0.9	764	21	ABBR18282	Plasmodium falcipa
86	14	1.1	693	24	AAE30108	Drosophila melanog	159	12	0.9	803	22	ABBS8104	Drosophila melanog
87	14	1.1	700	22	ABBS8943	Drosophila melanog	160	12	0.9	842	17	AAAM0137	Drosophila FRU 2 p
88	14	1.1	730	21	AAAY9949	Xenopus laevis kin	161	12	0.9	995	22	ABBS2510	Drosophila melanog
89	14	1.1	955	15	AAAR57365	K39 polypeptide of	162	12	0.9	1023	22	ABBR11436	D. discoidem ster
90	14	1.1	955	17	AAAR57365	Leishmania chagasi	163	12	0.9	1341	23	ABBR25703	Fungal ZBC protein
91	14	1.1	1245	21	AAAB18244	Plasmodium falcipa	164	12	0.9	1493	22	ABBS69228	Drosophila melanog
92	14	1.1	1460	21	AAAG40077	Arabidopsis thalia	165	12	0.9	1501	23	AAU76762	Plasmodium falcipa
93	14	1.1	1462	21	AAAG40076	Arabidopsis thalia	166	12	0.9	1516	21	AAAB18195	Plasmodium falcipa
94	14	1.1	1518	21	AAAG40075	Arabidopsis thalia	167	12	0.9	1568	24	ABBP70152	Amino acid sequenc
95	14	1.1	1604	21	AAAG31284	Arabidopsis thalia	168	12	0.9	2061	22	ABBS6938	Drosophila melanog
96	14	1.1	1606	21	AAAG31283	Arabidopsis thalia	169	12	0.9	3502	22	ABBS8382	Drosophila melanog
97	14	1.1	1662	21	AAAG31282	Arabidopsis thalia	170	12	0.9	5533	22	ABBS65772	Drosophila melanog
98	14	1.1	2010	21	AAAB18218	Plasmodium falcipa	171	12	0.9	5560	22	ABBR11160	Drosophila melanog
99	13	1.0	39	12	AAAR12902	CSP-2 peptide from	172	11	0.9	121	21	AAAG34475	Arabidopsis thalia
100	13	1.0	48	12	AAAR12903	CSP-2 peptide from	173	11	0.9	133	23	ABBP1443	Human ORFX protein
101	13	1.0	108	22	AAAO0353	Human polypeptide	174	11	0.9	137	22	AAO05799	Human polypeptide
102	13	1.0	138	21	AAAG11796	Arabidopsis thalia	175	11	0.9	160	22	AAAG74164	Human colon cancer
103	13	1.0	147	21	AAAG11795	Arabidopsis thalia	176	11	0.9	205	23	AAE14609	Human microtubule
104	13	1.0	282	21	AAAG35584	Arabidopsis thalia	177	11	0.9	205	24	ABU08213	Human kinesin moto
105	13	1.0	341	23	ABBB80079	Human kinesin moto	178	11	0.9	250	23	AAU93167	Arabidopsis transc
106	13	1.0	341	24	ABG72398	Human kinesin prot	179	11	0.9	278	22	ABBS6939	Drosophila melanog
107	13	1.0	363	21	AAAG35583	Arabidopsis thalia	180	11	0.9	306	22	ABBR70631	Drosophila melanog
108	13	1.0	367	21	AAAG35582	Arabidopsis thalia	181	11	0.9	314	23	AAU82970	C. albicans MTW1 p
109	13	1.0	397	23	ABP73400	Candida albicans e	182	11	0.9	338	23	ABBR07411	Human Hsk1p3A moto
110	13	1.0	507	22	ABG05000	Novel human diagno	183	11	0.9	343	23	AAU77185	Human kint-3 DNA f
111	13	1.0	679	22	ABBS8212	Drosophila melanog	184	11	0.9	348	22	ABG27541	Novel human diagno
112	13	1.0	790	21	AAAG11923	Arabidopsis thalia	185	11	0.9	361	23	ABBR08099	Human HSET motor p
113	13	1.0	793	22	ABBR1688	Drosophila melanog	186	11	0.9	363	23	AAU77186	Human kint-3 DNA f
114	13	1.0	794	21	AAAG41922	Arabidopsis thalia	187	11	0.9	369	23	ABBR08101	Human HSET motor p
115	13	1.0	814	21	AAAG41921	Arabidopsis thalia	188	11	0.9	370	23	ABBR08100	Human HSET motor p
116	13	1.0	941	23	ABP74093	Candida albicans e	189	11	0.9	373	23	AAU77183	Human kint-3 DNA f
117	13	1.0	1003	22	ABBB61405	Drosophila melanog	190	11	0.9	375	23	AAU79591	Human Hsk1f5, kin
118	13	1.0	1074	22	ABBB61124	Drosophila melanog	191	11	0.9	375	23	AAE14401	Human Hsk1f5, kin
119	13	1.0	1151	21	AAV83251	F-box protein Grt1	192	11	0.9	381	23	AAU77184	Human kint-3 DNA f
120	13	1.0	1151	23	AAE28254	Yeast Grt1 protein	193	11	0.9	383	23	ABBR07412	Amino acid sequenc
121	13	1.0	1212	22	ABBR71114	Drosophila melanog	194	11	0.9	383	24	ABG72053	Human Hsk1p3 fragm
122	13	1.0	1279	23	ABG70787	Human kinesin-rela	195	11	0.9	409	22	AAU15082	Protein encoded by
123	13	1.0	1279	23	ABBB80078	Human kinesin moto	196	11	0.9	409	23	ABBP73217	Candida albicans e
124	13	1.0	1279	24	ABG72397	Human partial kine	197	11	0.9	409	23	AAU79592	Human Hsk1f5, kin
125	13	1.0	1893	21	ABBS9829	Drosophila melanog	198	11	0.9	409	23	AAE14402	Human Hsk1f5 prot
126	13	1.0	2573	21	AAAB18234	Plasmodium falcipa	199	11	0.9	461	16	AAAR75506	Nicotiana alata ar
127	12	0.9	40	12	AAAR12904	CSP-2 peptide from	200	11	0.9	473	23	ABG70992	Human target prote
128	12	0.9	212	23	ABG93169	S. cerevisiae BAX-	201	11	0.9	522	23	ABG70991	Human Hsk1f4 const
129	12	0.9	246	22	AAAB47285	rcp41. Cryptospor	202	11	0.9	526	22	ABBS68634	Drosophila melanog
130	12	0.9	263	22	ABBR70399	Drosophila melanog	203	11	0.9	561	22	ABBS67878	Drosophila melanog
131	12	0.9	306	22	ABBB60719	Drosophila melanog	204	11	0.9	589	22	ABBS8929	Drosophila melanog
132	12	0.9	306	22	ABBB60722	Drosophila melanog	205	11	0.9	601	22	ABBB61643	Drosophila melanog
133	12	0.9	362	23	AAE14501	Human mitotic cent	206	11	0.9	673	22	AAAM38739	Human polypeptide
134	12	0.9	381	23	ABP73197	Candida albicans e	207	11	0.9	676	23	ABG93214	S. cerevisiae BAX-
135	12	0.9	391	23	AAE14500	Human mitotic cent	208	11	0.9	722	21	AAAB18291	Plasmodium falcipa
136	12	0.9	405	23	AAE14504	Human mitotic cent	209	11	0.9	774	22	ABBB64148	Drosophila melanog
137	12	0.9	410	21	ABBS6496	Human prostate can	210	11	0.9	776	22	AAAM40525	Human polypeptide
138	12	0.9	427	21	ABBS8461	Drosophila melanog	211	11	0.9	782	22	ABBR11348	Human kinesin rela
139	12	0.9	427	22	ABBS5442	Drosophila melanog	212	11	0.9	864	24	ABBR07410	Human kinesin moto
140	12	0.9	430	23	AAAB14499	Human mitotic cent	213	11	0.9	864	23	ABG72052	Human kinesin moto
141	12	0.9	434	23	AAE14503	Human mitotic cent	214	11	0.9	1025	23	ABBS2133	Plasmodium falcipa
142	12	0.9	473	23	AAE14502	Human mitotic cent	215	11	0.9	1029	22	ABBS20649	Novel human diagno
143	12	0.9	511	23	ABP73951	Candida albicans e	216	11	0.9	1122	21	AAAB18295	Plasmodium falcipa
144	12	0.9	638	22	ABBR71320	Drosophila melanog	217	11	0.9	1174	22	AAAB61704	Drosophila melanog
145	12	0.9	647	11	AAAR06001	Insect receptor po	218	11	0.9	1230	19	AAAB48895	Candida albicans C
146	12	0.9	675	17	AAAM00136	Drosophila FRU 1 p	219	11	0.9	1232	23	ABG70990	Human Hsk1f4 prote
147	12	0.9	679	20	ABBB8456	Human kinesin-rela	220	11	0.9	1234	23	ABG70993	Human Hsk1f4b prot
148	12	0.9	690	22	AAAM9257	Human polypeptide	221	11	0.9	1269	21	AAU77955	A. thaliana enviro
149	12	0.9	691	21	AAV32253	Drosophila diseati	222	11	0.9	1351	21	AAAB18290	Plasmodium falcipa
150	12	0.9	691	22	ABBB64167	Drosophila melanog	223	11	0.9	1368	23	AAU77182	Human kinesin moto
151	12	0.9	694	22	AAAM1043	Human polypeptide	224	11	0.9	1373	22	ABBR71036	Drosophila melanog
152	12	0.9	694	22	AAAM1044	Human polypeptide	225	11	0.9	1388	23	AAU79590	Human kinesin moto
153	12	0.9	723	23	AAE14505	Human mitotic cent	226	11	0.9	1388	23	AAE14400	Human kinesin supe
154	12	0.9	725	23	AAE14506	Human mitotic cent	227	11	0.9	1388	22	ABR48222	Human bladder can
155	12	0.9	728	23	AAAM39258	Human polypeptide	228	11	0.9	1848	24	ABBB68996	Drosophila melanog

229	11	0.9	1945	22	ABB64947	Drosophila melanog	302	10	0.8	633	21	AAG31266	Arabidopsis thaliana
230	11	0.9	1979	21	AA18171	Plasmodium falcipa	303	10	0.8	633	22	AAE01905	Arabidopsis thaliana
231	11	0.9	1997	21	AA18287	Plasmodium falcipa	304	10	0.8	644	22	ABE59973	Drosophila melanog
232	11	0.9	2088	22	ABB60656	Drosophila melanog	305	10	0.8	645	22	ABB63654	Drosophila melanog
233	11	0.9	3973	21	AA18253	Plasmodium falcipa	306	10	0.8	659	17	AAW09294	L-proline-4-hydrox
234	10	0.8	12	24	ABU14073	Transdominant elite	307	10	0.8	676	22	ABB63098	Drosophila melanog
235	10	0.8	13	23	ABP72103	Beta 1,2-N-acetylgl	308	10	0.8	679	22	AA45693	P. falciparum gp6E
236	10	0.8	16	22	AAU29011	Vector pMAL-P2 mal	309	10	0.8	690	21	AA456694	MBP-VT-1 subunit A
237	10	0.8	16	22	AA197228	Maltose binding pr	310	10	0.8	690	23	AAU77830	MBP-VT-1 fusion p
238	10	0.8	24	21	AA10777	B. subtilis lumazi	311	10	0.8	694	21	AA456695	MBP-VT-2 subunit A
239	10	0.8	24	22	AA170814	MBP/BMP fusion con	312	10	0.8	694	23	AAU77831	MBP-VT-2-A fusion p
240	10	0.8	26	22	AA170815	MBP/BMP fusion con	313	10	0.8	700	23	AA192360	Herbicide-inducible
241	10	0.8	30	22	AA170816	MBP/BMP fusion con	314	10	0.8	708	17	AAW06411	Maltose binding pr
242	10	0.8	40	18	AAW35993	Peptide SEQ ID NO:	315	10	0.8	708	21	AA456690	MBP-VT-1 subunit A
243	10	0.8	40	22	AAU26046	Peptide sequence o	316	10	0.8	708	23	AAU77826	MBP-VT-1-A fusion p
244	10	0.8	62	22	AAO01380	Human polypeptide	317	10	0.8	711	17	AAW06412	Maltose binding pr
245	10	0.8	67	21	AA456908	Peptide MSI-1922.	318	10	0.8	711	21	AA456691	MBP-VT-2 subunit A
246	10	0.8	91	10	AA190415	Plasmodium falcipa	319	10	0.8	711	21	AAU77827	MBP-VT-2-A fusion p
247	10	0.8	100	20	AAW73960	Human TP2/MBP fus	320	10	0.8	718	22	ABB66061	Drosophila melanog
248	10	0.8	100	23	AA151569	Maltose binding pr	321	10	0.8	720	22	AAE01346	Arabidopsis thaliana
249	10	0.8	108	22	AAO02887	Human polypeptide	322	10	0.8	732	22	ABB65450	Drosophila melanog
250	10	0.8	110	22	ABB64832	Drosophila melanog	323	10	0.8	740	23	ABG71293	Human oestrogen re
251	10	0.8	121	22	AA664025	Human polypeptide	324	10	0.8	740	23	AAW51591	Fused ER protein.
252	10	0.8	147	21	AA602949	Human secreted pro	325	10	0.8	750	22	AA190813	Amino acid sequenc
253	10	0.8	224	21	ABG10926	Novel human diagno	326	10	0.8	752	21	AAW90883	P. falciparum gp6E
254	10	0.8	240	23	ABG49146	Lieteria monocyclo	327	10	0.8	757	22	AAU19569	Human diagnostic a
255	10	0.8	248	22	ABG10925	Novel human diagno	328	10	0.8	757	23	ABP51294	Human MDOT SEQ ID
256	10	0.8	265	23	ABP73329	Candida albicans e	329	10	0.8	762	23	ABG60124	Human DITP polype
257	10	0.8	274	23	ABG70881	Human kinesin HsKi	330	10	0.8	763	21	AA182155	Pseudomonas putida
258	10	0.8	274	23	ABG79895	Human kinesin moto	331	10	0.8	784	20	AA182155	Thermomyces lanugi
259	10	0.8	274	23	ABG32861	Human kinesin HsKi	332	10	0.8	801	22	ABB62287	Drosophila melanog
260	10	0.8	281	22	ABG17932	Novel human diagno	333	10	0.8	809	23	AAW48930	Maltose binding pr
261	10	0.8	314	22	ABG14144	Novel human diagno	334	10	0.8	839	23	ABG71292	Human androgen rec
262	10	0.8	323	22	ABG60025	Drosophila melanog	335	10	0.8	839	23	AA668238	Fused androgen rec
263	10	0.8	329	23	AAE17787	Human kinesin supe	336	10	0.8	931	22	ABB61093	Drosophila melanog
264	10	0.8	355	22	ABB66677	Drosophila melanog	337	10	0.8	953	21	AA4582156	Pseudomonas putida
265	10	0.8	359	23	ABB84482	Human HsKi16b mot	338	10	0.8	988	22	ABB60007	Drosophila melanog
266	10	0.8	359	23	ABB79530	Human kinesin moto	339	10	0.8	1029	23	AAE17786	Human kinesin supe
267	10	0.8	359	23	AAE22526	Human HsKi16b mot	340	10	0.8	1038	23	AAW48337	Human kinesin supe
268	10	0.8	368	23	ABG93036	S. cerevisiae BAX-	341	10	0.8	1059	22	ABB58331	Drosophila melanog
269	10	0.8	378	22	ABB68211	Drosophila melanog	342	10	0.8	1068	23	ABP35714	Fungal Z8C protein
270	10	0.8	412	21	AA190266	Protein chimera Q.	343	10	0.8	1084	19	AAW59033	B. pallidus DNA po
271	10	0.8	412	24	ABG73864	L. infantum antige	344	10	0.8	1121	22	AB172021	Drosophila melanog
272	10	0.8	437	22	ABB62782	Drosophila melanog	345	10	0.8	1122	23	ABG93308	C. albicans BAX-as
273	10	0.8	437	22	ABB65444	Drosophila melanog	346	10	0.8	1198	22	ABB60371	Drosophila melanog
274	10	0.8	456	22	AA197389	Rice photosensitiv	347	10	0.8	1276	19	AAW58034	B. pallidus DNA po
275	10	0.8	468	22	AA197390	Rice photosensitiv	348	10	0.8	1308	21	AA181817	Plasmodium falcipa
276	10	0.8	488	23	ABP73873	Candida albicans e	349	10	0.8	1362	23	AAU74840	Human HsKi16b pr
277	10	0.8	500	19	AAW64057	Human IL-9 recepto	350	10	0.8	1375	23	ABB84481	Human HsKi16b pr
278	10	0.8	501	19	AAW64056	Human IL-9 recepto	351	10	0.8	1375	23	ABB79531	Human kinesin moto
279	10	0.8	501	19	AAW64055	Human IL-9 recepto	352	10	0.8	1375	23	AAE22525	Human HsKi16b pr
280	10	0.8	503	21	AA1863190	Human secreted pro	353	10	0.8	1392	24	AAE3129	Human cytoskeleton
281	10	0.8	504	21	AA1863189	Gene 5 human secre	354	10	0.8	1419	23	ABP52134	Plasmodium falcipa
282	10	0.8	504	22	ABB63967	Drosophila melanog	355	10	0.8	1438	21	AA182270	Plasmodium falcipa
283	10	0.8	511	23	AAW51137	Maltose binding pr	356	10	0.8	1598	23	ABP73929	Arabidopsis thaliana
284	10	0.8	514	21	AAW31268	Arabidopsis thaliana	357	10	0.8	1805	23	ABP68930	Candida albicans e
285	10	0.8	515	21	AAW31267	Arabidopsis thaliana	358	10	0.8	1817	21	AA181825	Human polypeptide
286	10	0.8	521	24	AAO19700	IL-13 modulation m	359	10	0.8	1838	22	ABB64225	Plasmodium falcipa
287	10	0.8	522	24	ABB81964	Human IL-9 recepto	360	10	0.8	1921	22	ABB62962	Drosophila melanog
288	10	0.8	530	22	ABB61485	Drosophila melanog	361	10	0.8	1926	22	AAW484915	Drosophila melanog
289	10	0.8	535	16	AA178524	Maltose binding pr	362	10	0.8	1991	22	ABB60651	Shrimp white spot
290	10	0.8	537	7	AA178524	Maltose binding pr	363	10	0.8	2112	22	ABB60403	Drosophila melanog
291	10	0.8	542	23	AAU74356	Sequence of the As	364	10	0.8	2183	22	ABG29838	Novel human diagno
292	10	0.8	544	23	ABP74012	Candida albicans e	365	10	0.8	2435	22	ABB60448	Drosophila melanog
293	10	0.8	549	22	ABB66830	Drosophila melanog	366	10	0.8	2532	22	ABB63465	Drosophila melanog
294	10	0.8	573	21	AA182154	Pseudomonas putida	367	10	0.8	2532	22	ABB61616	Drosophila melanog
295	10	0.8	575	22	ABB63518	Drosophila melanog	368	10	0.8	3614	20	ABB62664	Drosophila melanog
296	10	0.8	580	21	AA182157	Pseudomonas putida	369	10	0.8	4134	20	AA191946	Plasmodium falcipa
297	10	0.8	580	22	ABB62524	Drosophila melanog	370	10	0.8	4904	22	ABB62045	Drosophila melanog
298	10	0.8	596	23	ABG68928	Plasmodium falcipa	371	9	0.7	34	23	ABG40545	Human peptide enco
299	10	0.8	609	22	ABB61332	Drosophila melanog	372	9	0.7	39	18	AAW33332	Sequence encoded b
300	10	0.8	610	22	ABP71317	Drosophila melanog	373	9	0.7	52	23	ABP31745	Human structural p
301	10	0.8	624	22	ABB70398	Drosophila melanog	374	9	0.7	57	22	ABG52907	Human liver peptid

375	9	0.7	57	22	ABR38053	448	9	0.7	856	22	ABB68808	Drosophila melanog
376	9	0.7	57	22	AAM5868	449	9	0.7	858	22	ABB60597	Drosophila melanog
377	9	0.7	57	22	AAM7195	450	9	0.7	860	19	AAM63700	Receptor type tyro
378	9	0.7	57	22	AAM31476	451	9	0.7	885	22	ABB60372	Drosophila melanog
379	9	0.7	57	22	ABG40992	452	9	0.7	887	22	ABB71107	Drosophila melanog
380	9	0.7	72	23	ABR69713	453	9	0.7	898	22	ABU53123	Intracellular traf
381	9	0.7	139	22	ABR69010	454	9	0.7	898	23	AAU76957	Novel human kinase
382	9	0.7	140	21	ACG09377	455	9	0.7	905	23	AAU76967	Novel human kinase
383	9	0.7	145	22	ABR71063	456	9	0.7	943	22	ABB62085	Drosophila melanog
384	9	0.7	159	21	ABA43261	457	9	0.7	950	22	ABR71271	Drosophila melanog
385	9	0.7	167	21	ACG31470	458	9	0.7	960	22	ABR58434	Drosophila melanog
386	9	0.7	176	24	ABR30194	459	9	0.7	1003	18	AAM19917	Drosophila vitellin
387	9	0.7	177	20	AAV59708	460	9	0.7	1018	22	ABB63718	Drosophila melanog
388	9	0.7	180	24	ACG30195	461	9	0.7	1036	22	ACG85011	Shrimp white spot
389	9	0.7	187	21	AAV45405	462	9	0.7	1094	23	ABP5676	Fungal ZBC protein
390	9	0.7	208	22	ABG15503	463	9	0.7	1140	22	ABR71145	Drosophila melanog
391	9	0.7	213	22	AAU19360	464	9	0.7	1143	23	ABP73852	Candida albicans e
392	9	0.7	248	21	AAV60582	465	9	0.7	1165	23	AAE24372	Fruit fly E93 prog
393	9	0.7	257	21	AAAB18340	466	9	0.7	1186	22	ABB63516	Drosophila melanog
394	9	0.7	272	21	AAV97188	467	9	0.7	1193	21	ABR18306	Plasmodium falcipa
395	9	0.7	280	22	ABR65783	468	9	0.7	1195	22	ABB63694	Drosophila melanog
396	9	0.7	284	22	ABR69768	469	9	0.7	1207	22	ABB62248	Drosophila melanog
397	9	0.7	285	22	ABR65204	470	9	0.7	1221	23	AAE24373	Fruit fly E93 prog
398	9	0.7	300	21	AAV31469	471	9	0.7	1245	23	ABP73820	Candida albicans e
399	9	0.7	313	22	ABR62420	472	9	0.7	1418	23	ABG93257	C. albicans BAX-as
400	9	0.7	314	22	ABR67734	473	9	0.7	1419	22	ABR65521	Drosophila melanog
401	9	0.7	320	21	AAV45404	474	9	0.7	1436	21	ABR18199	Plasmodium falcipa
402	9	0.7	323	21	AAV45403	475	9	0.7	1445	23	ABP75623	Fungal ZBC protein
403	9	0.7	324	24	ABJ19039	476	9	0.7	1460	22	ABR71977	Drosophila melanog
404	9	0.7	330	23	ABP9157	477	9	0.7	1476	22	ABR58706	Drosophila melanog
405	9	0.7	354	23	ABP73242	478	9	0.7	1518	22	ABR64629	Drosophila melanog
406	9	0.7	357	23	AAU76958	479	9	0.7	1537	22	ABR58058	Drosophila melanog
407	9	0.7	372	23	ABR93414	480	9	0.7	1557	22	ABR62485	Drosophila melanog
408	9	0.7	378	22	ABR60643	481	9	0.7	1557	23	ABG70015	Larval viability a
409	9	0.7	398	21	AAV29336	482	9	0.7	1714	21	ABR18275	Plasmodium falcipa
410	9	0.7	410	23	ABR55031	483	9	0.7	1905	22	ABR64925	Drosophila melanog
411	9	0.7	412	21	AAV97189	484	9	0.7	2037	22	ABR69099	Drosophila melanog
412	9	0.7	437	22	AAV45424	485	9	0.7	2206	21	ABR18254	Plasmodium falcipa
413	9	0.7	437	22	AAV62618	486	9	0.7	2485	21	ABR18172	Plasmodium falcipa
414	9	0.7	454	22	ABR30816	487	9	0.7	2938	15	AAV59923	GAP protein Irai.
415	9	0.7	454	22	ABP73979	488	9	0.7	4498	22	ABR58595	Drosophila melanog
416	9	0.7	470	22	ABR59219	489	9	0.6	10	22	AAV83925	Arabidopsis thalia
417	9	0.7	498	23	ABP73393	490	9	0.6	27	17	AAV94328	Anti-thrombogenic
418	9	0.7	502	22	ABR65558	491	9	0.6	39	21	AAV58570	Arabidopsis thalia
419	9	0.7	508	21	ABR18187	492	9	0.6	39	22	ABG58895	Human liver peptid
420	9	0.7	525	22	ABR67142	493	9	0.6	39	22	ABR43516	Peptide #11022 enc
421	9	0.7	548	22	ABR59009	494	9	0.6	39	22	ABR26476	Human brain expres
422	9	0.7	567	21	ABR18197	495	9	0.6	39	22	ABR64453	Human brain expres
423	9	0.7	568	22	ABR58131	496	9	0.6	39	22	AAV77259	Human bone marrow
424	9	0.7	578	22	ABR59076	497	9	0.6	39	22	AAV21195	Peptide #7629 enc
425	9	0.7	604	23	ABR05593	498	9	0.6	39	22	AAV37410	Peptide #11447 enc
426	9	0.7	613	22	ABR65716	499	9	0.6	39	23	ABG46275	Human peptid enc
427	9	0.7	621	22	ABR58165	500	9	0.6	41	22	ABG55641	Human liver peptid
428	9	0.7	629	22	AAV70887	501	9	0.6	41	22	ABR40366	Peptide #7872 enc
429	9	0.7	645	22	ABR61545	502	9	0.6	41	22	AAV73890	Human brain expres
430	9	0.7	648	22	ABR63109	503	9	0.6	41	22	AAV61175	Human bone marrow
431	9	0.7	660	22	ABR68167	504	9	0.6	41	22	AAV34072	Peptide #8109 enc
432	9	0.7	670	23	ABP73731	505	9	0.6	41	23	ABG43778	Human peptid enc
433	9	0.7	675	24	ABP78737	506	9	0.6	42	22	ABR68447	Drosophila melanog
434	9	0.7	691	22	ABR57876	507	9	0.6	52	22	ABG48060	Human liver peptid
435	9	0.7	719	22	AAV70700	508	9	0.6	52	22	ABR28036	Human peptid #687
436	9	0.7	724	22	ABR65778	509	9	0.6	52	22	ABR33210	Peptide #716 enc
437	9	0.7	726	21	ACG31118	510	9	0.6	52	22	ABR18675	Protein #674 enc
438	9	0.7	729	22	AAV70847	511	9	0.6	52	22	AAV54003	Human brain expres
439	9	0.7	740	23	ABP73836	512	9	0.6	52	22	AAV66393	Human bone marrow
440	9	0.7	764	23	ABP73760	513	9	0.6	52	22	AAV14263	Peptide #697 enc
441	9	0.7	769	23	ABP73280	514	9	0.6	52	22	AAV26673	Peptide #710 enc
442	9	0.7	807	21	AAV18311	515	9	0.6	52	22	AAV01993	Peptide #675 enc
443	9	0.7	816	24	AAV30116	516	9	0.6	52	23	ABG36042	Human peptid enc
444	9	0.7	829	21	ACG31117	517	9	0.6	63	21	AAV54171	Arabidopsis thalia
445	9	0.7	834	21	AAV31116	518	9	0.6	71	22	ABR70696	Drosophila melanog
446	9	0.7	844	22	ABR66588	519	9	0.6	75	22	ABG55613	Human liver peptid
447	9	0.7	856	22	ABR60983	520	9	0.6	75	22	AAV61148	Human brain expres

521	8	0.6	75	22	AAW73862	Human bone marrow
522	8	0.6	75	23	ABG43750	Human peptide enco
523	8	0.6	76	19	AAW72051	HSV-2 strain SB5 C
524	8	0.6	82	21	AAH11545	SEN virus protein
525	8	0.6	89	23	ABP31402	Human ORF375 prote
526	8	0.6	90	18	AAH38201	Rice SCARECROW SRP
527	8	0.6	90	21	AAH28581	Rice SCL4. Oryza
528	8	0.6	95	22	ABG26592	Novel human diago
529	8	0.6	102	22	ABG26896	Novel human diago
530	8	0.6	103	22	ABG28378	Novel human diago
531	8	0.6	107	22	AAO07842	Human polypeptide
532	8	0.6	114	22	ABH11091	Human nuclear pore
533	8	0.6	117	22	ABH67854	Drosophila melanog
534	8	0.6	117	22	ABG27133	Novel human diago
535	8	0.6	126	22	ABG03470	Novel human diago
536	8	0.6	134	21	AAG29846	Arbidiopsis thalia
537	8	0.6	134	22	AAO01002	Human polypeptide
538	8	0.6	134	22	AAO12584	Human polypeptide
539	8	0.6	134	24	ABP81236	Arbidiopsis thalia
540	8	0.6	137	23	ABJ11269	Yeast selected int
541	8	0.6	137	23	ABG77367	Selected interacti
542	8	0.6	141	21	AAG41855	Arbidiopsis thalia
543	8	0.6	143	22	ABP71001	Drosophila melanog
544	8	0.6	145	21	AAG47273	Arbidiopsis thalia
545	8	0.6	146	21	AAI96215	Argemone mexicana
546	8	0.6	146	24	ABR40724	Arbidiopsis thalia
547	8	0.6	147	21	AAI96215	Argemone mexicana
548	8	0.6	147	23	ABH83125	Human glutary-CoA
549	8	0.6	151	21	AAI29944	Arbidiopsis thalia
550	8	0.6	154	22	ABH70370	Drosophila melanog
551	8	0.6	157	21	AAI29944	Arbidiopsis thalia
552	8	0.6	164	19	AAH80406	A secreted protein
553	8	0.6	164	23	ABP61824	Human polypeptide
554	8	0.6	167	22	ABH62385	Drosophila melanog
555	8	0.6	173	21	AAI08442	Arbidiopsis thalia
556	8	0.6	173	21	AAI08442	Arbidiopsis thalia
557	8	0.6	184	21	AAI08442	Arbidiopsis thalia
558	8	0.6	185	23	ABJ11066	Yeast selected int
559	8	0.6	185	23	ABG77195	Selected interacti
560	8	0.6	186	22	ABH69467	Drosophila melanog
561	8	0.6	193	24	ABP71417	Human UBE2 related
562	8	0.6	195	19	AAH59053	His rfcF10 protein
563	8	0.6	215	18	AAW24049	Rat fibroblast gro
564	8	0.6	216	23	ABJ11078	Yeast selected int
565	8	0.6	216	23	ABG77207	Selected interacti
566	8	0.6	221	22	ABH64291	Drosophila melanog
567	8	0.6	221	23	ABH06719	Human oestrogen re
568	8	0.6	222	23	ABP73512	Candida albicans e
569	8	0.6	227	21	AAI51369	Arbidiopsis thalia
570	8	0.6	228	21	AAI51369	Arbidiopsis thalia
571	8	0.6	228	22	ABH62497	Drosophila melanog
572	8	0.6	229	21	AAH43838	Human cancer assoc
573	8	0.6	229	21	AAI17420	Arbidiopsis thalia
574	8	0.6	232	21	AAI57272	Arbidiopsis thalia
575	8	0.6	234	21	AAI51368	Arbidiopsis thalia
576	8	0.6	234	23	ABJ11125	Yeast selected int
577	8	0.6	234	23	ABG77254	Selected interacti
578	8	0.6	235	21	AAI51368	Arbidiopsis thalia
579	8	0.6	239	21	AAI57271	Arbidiopsis thalia
580	8	0.6	240	22	ABH70127	Drosophila melanog
581	8	0.6	242	22	ABH69486	Drosophila melanog
582	8	0.6	249	22	ABH60887	Drosophila melanog
583	8	0.6	250	24	AAU93090	Arbidiopsis thalia
584	8	0.6	250	24	ABR40794	Arbidiopsis thalia
585	8	0.6	251	23	ABP73583	Candida albicans e
586	8	0.6	255	21	AAI29944	Arbidiopsis thalia
587	8	0.6	255	21	AAI29944	Arbidiopsis thalia
588	8	0.6	258	21	AAI29944	Arbidiopsis thalia
589	8	0.6	259	21	AAI29944	Arbidiopsis thalia
590	8	0.6	259	21	AAI29944	Arbidiopsis thalia
591	8	0.6	260	21	AAI29944	Arbidiopsis thalia
592	8	0.6	260	21	AAI29944	Arbidiopsis thalia
593	8	0.6	265	22	ABG23353	Novel human diago

667	8	0.6	443	22	ABG26895	Novel human diagno	740	8	0.6	787	23	ABE08247	Human HsKlfC2. Ho
668	8	0.6	444	19	AAW60579	Drosophila neuron-	741	8	0.6	797	20	AAV05850	Banana ripening fr
669	8	0.6	445	21	AAAB0060	VRN2 polypeptide.	742	8	0.6	798	22	ABE62356	Drosophila melanog
670	8	0.6	446	19	AAW69563	Sugarbeet choline	743	8	0.6	809	23	ABP73854	Candida albicans e
671	8	0.6	448	22	ABE58601	Drosophila melanog	744	8	0.6	811	22	ABG23352	Novel human diagno
672	8	0.6	461	12	AAK14150	MSP encoded by pUC	745	8	0.6	812	21	AAV95985	Arabidopsis GABA r
673	8	0.6	461	12	AAK14530	Usp45 protein. La	746	8	0.6	814	22	ABE65317	Drosophila melanog
674	8	0.6	464	23	ABE93400	Herbically activ	747	8	0.6	825	22	ABE59301	Drosophila melanog
675	8	0.6	464	23	AAW47592	Drosophila cell cy	748	8	0.6	850	22	ABE65764	Drosophila melanog
676	8	0.6	469	22	ABE61392	Drosophila melanog	749	8	0.6	854	22	AAE67391	Partial amino acid
677	8	0.6	478	21	ABAB0728	Human IL-2/ P. yoe	750	8	0.6	865	23	ABP35641	Fungal ZBC protein
678	8	0.6	489	22	ABAB70861	Arabidopsis melanog	751	8	0.6	865	23	ABP35642	Fungal ZBC protein
679	8	0.6	497	21	AAAG0660	Arabidopsis thalia	752	8	0.6	875	19	AAW34987	Banika goldi endo
680	8	0.6	498	21	AAAG54168	Arabidopsis thalia	753	8	0.6	880	22	AAW70938	S. pombe potassum
681	8	0.6	510	21	AAAG3697	Arabidopsis thalia	754	8	0.6	893	22	ABE64865	Drosophila melanog
682	8	0.6	525	23	ABG91725	Purine/pyrimidine	755	8	0.6	900	22	ABE60688	Drosophila melanog
683	8	0.6	528	23	ABP73734	Candida albicans e	756	8	0.6	916	22	ABE63615	Drosophila melanog
684	8	0.6	529	23	ABG77458	Selected interact	757	8	0.6	927	21	AAV95984	Arabidopsis GABA r
685	8	0.6	530	23	ABU11360	Yeast selected int	758	8	0.6	938	22	ABE60562	Drosophila melanog
686	8	0.6	531	22	ABE70528	Drosophila melanog	759	8	0.6	944	21	ABE43351	Human ORFX ORF3115
687	8	0.6	535	22	ABE58106	Drosophila melanog	760	8	0.6	945	23	ABP73277	Candida albicans e
688	8	0.6	535	22	ABE58915	Drosophila melanog	761	8	0.6	951	22	ABE67950	Drosophila melanog
689	8	0.6	535	22	ABE52665	Herbically activ	762	8	0.6	959	22	AAE64950	Human Creaml prote
690	8	0.6	537	23	AAO14997	Rhodotorula minuta	763	8	0.6	970	23	ABP35631	Fungal ZBC protein
691	8	0.6	540	22	AAAG4948	Shrimp white spot	764	8	0.6	976	22	ABE65498	Drosophila melanog
692	8	0.6	541	22	ABE68621	Drosophila melanog	765	8	0.6	984	9	AAE0550	Malaria-specific 1
693	8	0.6	545	22	ABE70472	Drosophila melanog	766	8	0.6	984	13	AAE22247	SERR antigen. Pia
694	8	0.6	554	19	AAW69223	HAP4 protein. Sac	767	8	0.6	984	13	AAE27529	Plasmodium falci
695	8	0.6	559	24	ABU02325	S. pneumoniae type	768	8	0.6	984	15	AAE60164	SERR protein encod
696	8	0.6	566	24	ABE40876	Arabidopsis thalia	769	8	0.6	984	16	AAE68837	Plasmodium falci
697	8	0.6	576	21	AAAG38679	Arabidopsis thalia	770	8	0.6	984	20	AAV29589	Plasmodium falci
698	8	0.6	581	23	ABP25491	Streptococcus poly	771	8	0.6	998	22	ABE60423	Drosophila melanog
699	8	0.6	585	11	AAAR07274	Reverse transcript	772	8	0.6	1003	22	ABE68762	Drosophila melanog
700	8	0.6	586	12	AAAR1920	Reverse transcript	773	8	0.6	1008	22	ABE59260	Drosophila melanog
701	8	0.6	586	16	AAAR73935	msDNA-Ec67 reverse	774	8	0.6	1015	23	ABP73325	Candida albicans e
702	8	0.6	586	21	AAV76811	E. coli msDNA-Ec67	775	8	0.6	1028	22	ABE62708	Drosophila melanog
703	8	0.6	586	23	AAW49402	Escherichia coli C	776	8	0.6	1032	21	AAE45673	Arabidopsis thalia
704	8	0.6	594	24	ABP81457	Streptococcus pneu	777	8	0.6	1037	22	ABE70288	Arabidopsis thalia
705	8	0.6	603	24	ABU11606	Human MDT polypep	778	8	0.6	1045	21	AAE45672	Arabidopsis thalia
706	8	0.6	605	22	AAE04789	Lycopodium escul	779	8	0.6	1058	22	AAE68341	A human dishevelle
707	8	0.6	605	22	AAE72308	Neoxanthin cleavag	780	8	0.6	1090	21	AAE45671	Arabidopsis thalia
708	8	0.6	627	20	AAV05853	Banana ripening fr	781	8	0.6	1108	22	ABE63413	Drosophila melanog
709	8	0.6	634	22	ABE71624	Drosophila melanog	782	8	0.6	1132	21	AAE32139	Arabidopsis thalia
710	8	0.6	650	22	AAW37933	Human polypeptide,	783	8	0.6	1152	21	AAE32138	Arabidopsis thalia
711	8	0.6	650	23	AAE68957	Human polypeptide	784	8	0.6	1157	22	ABE62647	Drosophila melanog
712	8	0.6	652	18	AAW18010	Plasmodium falci	785	8	0.6	1173	22	ABE62245	Drosophila melanog
713	8	0.6	657	22	ABE57783	Drosophila melanog	786	8	0.6	1182	22	ABE64438	Drosophila melanog
714	8	0.6	663	18	AAW29150	Dual-specific murt	787	8	0.6	1200	22	ABE63533	Drosophila melanog
715	8	0.6	667	22	ABE62659	Drosophila melanog	788	8	0.6	1205	20	AAV33683	P. falciptarum 1-de
716	8	0.6	668	22	ABE66214	Drosophila melanog	789	8	0.6	1205	21	AAW90882	Isoprenoid related
717	8	0.6	673	22	ABE59674	Drosophila melanog	790	8	0.6	1205	23	AAO21863	Drosophila melanog
718	8	0.6	677	22	ABE62562	Drosophila melanog	791	8	0.6	1238	22	ABE61259	Drosophila melanog
719	8	0.6	677	22	ABE62568	Drosophila melanog	792	8	0.6	1242	20	AAV13461	Amino acid sequenc
720	8	0.6	677	22	ABE65183	Drosophila melanog	793	8	0.6	1242	20	AAW93972	Human IRS-1 and IR
721	8	0.6	678	14	AAE42087	D. melanogaster do	794	8	0.6	1242	22	AAE83921	Amino acid sequenc
722	8	0.6	710	23	ABP73976	Candida albicans e	795	8	0.6	1242	24	ABE59797	Human Insulin-rece
723	8	0.6	713	12	AAAR13991	Plasmodium falci	796	8	0.6	1242	24	ABE72368	Human Insulin-rece
724	8	0.6	717	22	ABE71958	Drosophila melanog	797	8	0.6	1242	24	ABE72369	Human Insulin-rece
725	8	0.6	718	22	ABE66734	Drosophila melanog	798	8	0.6	1242	24	ABE72370	Human Insulin-rece
726	8	0.6	725	23	AAU93157	Arabidopsis tiran	799	8	0.6	1242	24	ABE72371	Human Insulin-rece
727	8	0.6	728	22	ABE64748	Drosophila melanog	800	8	0.6	1242	24	ABE72372	Human Insulin-rece
728	8	0.6	743	22	ABE68417	Drosophila melanog	801	8	0.6	1243	16	AAE67708	Insulin receptor s
729	8	0.6	745	22	ABE71016	Drosophila melanog	802	8	0.6	1247	21	AAE18215	Plasmodium falci
730	8	0.6	751	14	AAE33424	GEPD. Plasmodium	803	8	0.6	1278	21	AAE18277	Plasmodium falci
731	8	0.6	757	21	AAE41423	Arabidopsis thalia	804	8	0.6	1311	22	ABE61244	Drosophila melanog
732	8	0.6	765	22	ABE70099	Drosophila melanog	805	8	0.6	1316	22	ABE69283	Drosophila melanog
733	8	0.6	769	22	ABE58091	Drosophila melanog	806	8	0.6	1321	17	AAE69694	Mouse IRS-2. Mus
734	8	0.6	770	21	AAE50814	D. melanogaster ac	807	8	0.6	1398	21	AAE18292	Plasmodium falci
735	8	0.6	772	21	AAE41422	Arabidopsis thalia	808	8	0.6	1402	23	ABE66040	Bifidobacterium 1o
736	8	0.6	780	22	AAE42085	Human polypeptide	809	8	0.6	1406	22	ABE71963	Drosophila melanog
737	8	0.6	784	22	ABE71112	Drosophila melanog	810	8	0.6	1412	22	ABE60943	Drosophila melanog
738	8	0.6	787	22	ABE71302	Shrimp white spot	811	8	0.6	1430	22	AAE85029	Human polypeptide
739	8	0.6	787	22	AAU99335	Human kinesin moto	812	8	0.6	1475	22	AAE40232	Human polypeptide

813	8	0.6	1486	22	ABB59325	Drosophila melanog	886	7	0.5	14	22	AAE05023	Cryptosporidium pa
814	8	0.6	1499	22	AAW42018	Human polypeptide	887	7	0.5	16	23	ABG96233	Cysteine-containin
815	8	0.6	1629	22	ABB63122	Drosophila melanog	888	7	0.5	14	21	AAE08561	Peptide identified
816	8	0.6	1679	22	ABB60498	Drosophila melanog	889	7	0.5	16	22	AAE05034	Cryptosporidium pa
817	8	0.6	1679	22	ABB60502	Drosophila melanog	890	7	0.5	17	22	ABG59600	Human liver peptid
818	8	0.6	1703	22	ABB66223	Drosophila melanog	891	7	0.5	17	22	ABB44227	Peptide #1173 enc
819	8	0.6	1850	22	AAH72210	Modified Chicken v	892	7	0.5	17	22	ABB27103	Protein #9102 enc
820	8	0.6	1980	22	ABB67589	Drosophila melanog	893	7	0.5	17	22	AAE65259	Human brain expres
821	8	0.6	2023	22	ABB63487	Drosophila melanog	894	7	0.5	17	22	AAE77964	Human bone marrow
822	8	0.6	2070	22	ABG03703	Novel human diagno	895	7	0.5	17	22	AAE21858	Peptide #8292 enc
823	8	0.6	2175	22	ABB65698	Drosophila melanog	896	7	0.5	17	22	AAE38182	Peptide #12219 enc
824	8	0.6	2274	22	ABB58657	Drosophila melanog	897	7	0.5	17	22	ABB74469	DNA repair protein
825	8	0.6	2314	22	AAE69136	M. catarrhalis les	898	7	0.5	19	23	ABG51474	Human liver peptid
826	8	0.6	2391	15	AAE56994	Carbamoyl-phosphat	899	7	0.5	19	22	ABB31409	Peptide #4060 enc
827	8	0.6	2441	22	ABB62231	Drosophila melanog	900	7	0.5	19	22	ABB36623	Peptide #4129 enc
828	8	0.6	2500	21	AAH18272	Plasmodium falcipa	901	7	0.5	19	22	ABB21958	Protein #13957 enc
829	8	0.6	2543	22	ABG21295	Novel human diagno	902	7	0.5	19	22	AAE57384	Human brain expres
830	8	0.6	3005	22	ABB58102	Drosophila melanog	903	7	0.5	19	22	AAE69777	Human bone marrow
831	8	0.6	3190	22	AAE84634	Amino acid sequenc	904	7	0.5	19	22	AAE17602	Peptide #4036 enc
832	8	0.6	3332	22	ABB69873	Drosophila melanog	905	7	0.5	19	22	AAE30122	Peptide #4159 enc
833	8	0.6	3536	22	ABB65480	Drosophila melanog	906	7	0.5	19	22	AAE05262	Peptide #3944 enc
834	8	0.6	3726	22	ABB63947	Drosophila melanog	907	7	0.5	19	23	ABG38409	Human peptid enc
835	8	0.6	4152	22	AAE84992	Shrimp white spot	908	7	0.5	20	24	ABJ38224	Human cycomegalovi
836	8	0.6	6025	22	AAE84939	Shrimp white spot	909	7	0.5	22	18	AAW45350	Peptide #7 bearing
837	8	0.6	6077	22	ABP95996	White spot syndrom	910	7	0.5	22	22	ABG56539	Human liver peptid
838	7	0.5	7	21	AAE15678	Single-chain antib	911	7	0.5	22	22	ABE41096	Peptide #8602 enc
839	7	0.5	10	22	AAE43208	Mycoplasma genital	912	7	0.5	22	22	ABE25149	Protein #7148 enc
840	7	0.5	10	22	AAE43210	Mycoplasma genital	913	7	0.5	22	22	AAE61952	Human brain expres
841	7	0.5	10	22	AAE43212	Mycoplasma genital	914	7	0.5	22	22	AAE74756	Human bone marrow
842	7	0.5	10	22	AAE43214	Mycoplasma genital	915	7	0.5	22	22	AAE20402	Peptide #6836 enc
843	7	0.5	10	22	AAE43216	Mycoplasma genital	916	7	0.5	22	22	AAE34872	Peptide #8909 enc
844	7	0.5	10	22	AAE43218	Mycoplasma genital	917	7	0.5	22	23	AAE44553	Human peptid enc
845	7	0.5	10	22	AAE43220	Mycoplasma genital	918	7	0.5	23	23	AAE22579	Drosophila MBD-1ik
846	7	0.5	10	22	AAE43222	Mycoplasma genital	919	7	0.5	24	23	ABG69758	Polypeptide identi
847	7	0.5	10	22	AAE43224	Mycoplasma genital	920	7	0.5	25	22	ABG57335	Human liver peptid
848	7	0.5	10	22	AAE43226	Mycoplasma genital	921	7	0.5	25	22	ABE41900	Peptide #9406 enc
849	7	0.5	10	22	AAE43228	Mycoplasma genital	922	7	0.5	25	22	AAE62779	Human brain expres
850	7	0.5	10	22	AAE43230	Mycoplasma genital	923	7	0.5	25	22	AAE75595	Human bone marrow
851	7	0.5	10	22	AAE43232	Mycoplasma genital	924	7	0.5	25	22	AAE35702	Peptide #9739 enc
852	7	0.5	10	22	AAE43234	Mycoplasma genital	925	7	0.5	26	21	AAE09850	Endoglucanase frag
853	7	0.5	10	22	AAE43236	Mycoplasma genital	926	7	0.5	26	22	ABG63388	Human liver peptid
854	7	0.5	10	22	AAE43238	Mycoplasma genital	927	7	0.5	26	22	ABE40942	Peptide #8448 enc
855	7	0.5	10	22	AAE43240	Mycoplasma genital	928	7	0.5	26	22	AAE61804	Human brain expres
856	7	0.5	10	22	AAE43242	Mycoplasma genital	929	7	0.5	26	22	AAE74603	Human bone marrow
857	7	0.5	10	22	AAE43244	Mycoplasma genital	930	7	0.5	26	22	AAE34715	Peptide #8752 enc
858	7	0.5	10	22	AAE43246	Mycoplasma genital	931	7	0.5	30	22	ABG58601	Human liver peptid
859	7	0.5	10	22	AAE43248	Mycoplasma genital	932	7	0.5	30	22	ABG59523	Human liver peptid
860	7	0.5	10	22	AAE43250	Mycoplasma genital	933	7	0.5	30	22	ABE43208	Peptide #10714 enc
861	7	0.5	10	22	AAE43252	Mycoplasma genital	934	7	0.5	30	22	ABE44143	Peptide #11649 enc
862	7	0.5	10	22	AAE43256	Mycoplasma genital	935	7	0.5	30	22	ABE26297	Protein #8296 enc
863	7	0.5	10	22	AAE43258	Mycoplasma genital	936	7	0.5	30	22	ABE27027	Human brain expres
864	7	0.5	10	22	AAE83890	Arbidiopsis thaila	937	7	0.5	30	22	AAE64114	Human liver expres
865	7	0.5	10	22	AAE83892	Arbidiopsis thaila	938	7	0.5	30	22	AAE76940	Human bone marrow
866	7	0.5	10	22	AAE83894	Arbidiopsis thaila	939	7	0.5	30	22	AAE77875	Human bone marrow
867	7	0.5	10	22	AAE83896	Arbidiopsis thaila	940	7	0.5	30	22	AAE21033	Peptide #7467 enc
868	7	0.5	10	22	AAE83898	Arbidiopsis thaila	941	7	0.5	30	22	AAE21772	Peptide #8206 enc
869	7	0.5	10	22	AAE83904	Arbidiopsis thaila	942	7	0.5	30	22	AAE37048	Peptide #11085 enc
870	7	0.5	10	22	AAE83906	Arbidiopsis thaila	943	7	0.5	30	22	AAE38096	Peptide #12133 enc
871	7	0.5	10	22	AAE83908	Arbidiopsis thaila	944	7	0.5	30	22	ABE46049	Human peptid enc
872	7	0.5	10	22	AAE83910	Arbidiopsis thaila	945	7	0.5	30	22	ABE46899	Human peptid enc
873	7	0.5	10	22	AAE83912	Arbidiopsis thaila	946	7	0.5	31	22	AAU18263	Novel human DNA-bi
874	7	0.5	10	22	AAE83914	Arbidiopsis thaila	947	7	0.5	31	23	ABG92684	Human DNA-binding
875	7	0.5	10	22	AAE83923	Arbidiopsis thaila	948	7	0.5	32	22	ABG50161	Human liver peptid
876	7	0.5	10	22	AAE83927	Arbidiopsis thaila	949	7	0.5	32	22	ABE30116	Peptide #2767 enc
877	7	0.5	10	22	AAE84194	Arbidiopsis thaila	950	7	0.5	32	22	ABE35288	Peptide #2794 enc
878	7	0.5	10	22	AAE84198	Arbidiopsis thaila	951	7	0.5	32	22	ABE20731	Protein #2730 enc
879	7	0.5	10	22	AAE84368	Arbidiopsis thaila	952	7	0.5	32	22	AAE56117	Human brain expres
880	7	0.5	10	22	AAE84370	Arbidiopsis thaila	953	7	0.5	32	22	AAE68489	Human bone marrow
881	7	0.5	10	22	AAE87012	Saccharomyces cere	954	7	0.5	32	22	AAE16297	Peptide #2731 enc
882	7	0.5	10	22	AAE893544	Forthead related c	955	7	0.5	32	22	AAE28787	Peptide #2824 enc
883	7	0.5	10	22	AAE93545	Forthead related c	956	7	0.5	32	22	AAE04032	Peptide #2714 enc
884	7	0.5	10	22	AAE93546	Forthead related c	957	7	0.5	32	23	ABE38071	Human peptid enc
885	7	0.5	13	18	AAW45349	Peptide #6 Bearing	958	7	0.5	33	22	ABG50270	Human liver peptid

959	7	0.5	33	22	ABG58748	Human liver peptid
960	7	0.5	33	22	ABR310242	Peptide #12893
961	7	0.5	33	22	ABR35408	Peptide #2914
962	7	0.5	33	22	ABR20845	Protein #2844
963	7	0.5	33	22	AAW56232	Human brain expres
964	7	0.5	33	22	AAW68610	Human bone marrow
965	7	0.5	33	22	AAW16417	Peptide #2851
966	7	0.5	33	22	AAW28916	Peptide #2853
967	7	0.5	33	22	AAW04151	Peptide #2833
968	7	0.5	33	23	ABG38189	Human peptid
969	7	0.5	35	21	ABG11032	Arabidopsis thalia
970	7	0.5	36	21	AAG14995	Arabidopsis thalia
971	7	0.5	40	22	AAO04685	Human polypeptid
972	7	0.5	42	22	AAO04746	Human polypeptid
973	7	0.5	44	21	AAG47925	Arabidopsis thalia
974	7	0.5	44	22	ABG54345	Human liver peptid
975	7	0.5	44	22	ABR39366	Peptide #6872
976	7	0.5	44	22	AAW60064	Human brain expres
977	7	0.5	44	22	AAW72647	Human bone marrow
978	7	0.5	44	22	AAW32881	Peptide #6918
979	7	0.5	44	23	ABG42471	Human peptid
980	7	0.5	45	17	AAW08985	Lysine/serine conc
981	7	0.5	46	21	AAG57114	Arabidopsis thalia
982	7	0.5	46	22	ABG57063	Human liver peptid
983	7	0.5	46	22	ABR69029	Drosophila melan
984	7	0.5	46	22	ABR41622	Peptide #9128
985	7	0.5	46	22	AAW62493	Human brain expres
986	7	0.5	46	22	AAW75302	Human bone marrow
987	7	0.5	46	22	AAW35415	Peptide #9452
988	7	0.5	48	21	AAW08479	Arabidopsis thalia
989	7	0.5	48	21	AAW47962	Arabidopsis thalia
990	7	0.5	50	23	ABP10282	Human ORF protei
991	7	0.5	51	22	AAW07319	Human polypeptid
992	7	0.5	52	22	ABG53747	Human liver peptid
993	7	0.5	52	22	ABR38853	Peptide #6359
994	7	0.5	52	22	AAW59498	Human brain expres
995	7	0.5	52	22	AAW72063	Human bone marrow
996	7	0.5	52	23	AAW32338	Peptide #6365
997	7	0.5	52	22	ABG41878	Human peptid
998	7	0.5	53	22	ABG11771	Novel human diagn
999	7	0.5	53	22	ABG20393	Novel human diagn
1000	7	0.5	53	22	AAW02332	Human polypeptid

ALIGNMENTS

RESULT	1
AA041989	
ID	AA041989 standard; protein; 190 AA.
XX	AA041989;
AC	
DT	25-MAR-2003 (updated)
DT	21-APR-1994 (first entry)
XX	
DE	Staphylococcus enterotoxin B.
XX	
KW	Staphylococcus; enterotoxin B; SEB; super antigen; SAg; receptor;
KW	random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
KW	autoimmune disease.
XX	
OS	Staphylococcus aureus.
XX	
EH	Key
FT	Region
FT	Region
FT	Region
XX	
FN	W09314634-A1.
XX	
PD	05-AUG-1993.

XX	28-JAN-1993;	93MO-US00839.
PF		
XX	28-JAN-1992;	92US-0827540.
PR		
XX		
PA	(NAME-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.	
XX		
PI	Kappler JW, Marrack P;	
XX		
DR	WPI; 1993-320314/40.	
XX		
PT	Preventing or treating toxic effects of super antigens - by admin. of	
PT	new modified or mutated super antigen which induces antibodies but	
PT	not T-cell activation	
XX		
PS	Disclosure; Fig 3; 54pp; English.	
XX		
CC	This sequence represents the Staphylococcus enterotoxin B (SEB)	
CC	super antigen (SAg). Mutant versions of this sequence generated	
CC	by random mutation by PCR, may be used to modify the T-cell response	
CC	elicited by an antigen, by interacting with specific Vbeta elements	
CC	of the T-cell receptors. This T-cell response modulation may be	
CC	useful in a wide variety of autoimmune diseases, where self-reactive	
CC	T-cells may be activated by SAg that bind to particular Vbeta types.	
CC	(Updated on 25-MAR-2003 to correct PN field.)	
XX		
SQ	Sequence 190 AA;	
Query Match	1.3%; Score 17; DB 14; Length 190;	
Best Local Similarity	100.0%; Pred. No. 1,2e-07;	
Matches 17; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
OY	742 NNNNNNNNNNNNNNNNT 758	
DB	171 NNNNNNNNNNNNNNNNI 187	
RESULT 2		
AA41990		
ID	AA41990 standard; protein; 190 AA.	
XX		
AC	AA41990;	
XX		
DT	25-MAR-2003 (updated)	
DT	21-APR-1994 (first entry)	
XX		
DE	Staphylococcus enterotoxin B mutant BR-75.	
XX		
KM	Staphylococcus; enterotoxin B; SEB; super antigen; SAg; receptor;	
KW	random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;	
XX	autoimmune disease.	
OS		
XX	Staphylococcus aureus.	
XX		
Key	Location/Qualifiers	
FT	97..157	
FT	/note= "Region undisclosed in the specification"	
FT	162..186	
FT	/note= "Region undisclosed in the specification"	
XX		
PN	W09314634-A1.	
XX		
PD	05-AUG-1993.	
XX		
PF	28-JAN-1993; 93MO-US00839.	
XX		
PR	28-JAN-1992; 92US-0827540.	
XX		
PA	(NAME-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.	
XX		
PI	Kappler JW, Marrack P;	
XX		
DR	WPI; 1993-320314/40.	

XX Preventing or treating toxic effects of super antigens - by admin. of
PT new modified or mutated super antigen which induces antibodies but
PR not T-cell activation
XX
PS Disclosure; Fig 3; 54pp; English.
XX
CC The sequences given in AAR41990-2013 are mutant versions of the
CC Staphylococcus enterotoxin B (SEB) super antigen (SAG). These mutants
CC generated by random mutation by PCR, may be used to modify the T-cell
CC response elicited by an antigen, by interacting with specific Vbeta
CC elements of the T-cell receptors. This T-cell response modulation
CC may be useful in a wide variety of autoimmune diseases, where self-
CC reactive T-cells may be activated by SAG that bind to particular
CC Vbeta types.
CC (Updated on 25-MAR-2003 to correct PN field.)
CC
XX
SQ Sequence 190 AA;
XX
Query Match 1.3%; Score 17; DB 14; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 NNNNNNNNNNNNNNNNNI 758
DB 171 NNNNNNNNNNNNNNNNNI 187
XX
RESULT 3
AAR41991
ID AAR41991 standard; protein; 190 AA.
XX
AC AAR41991;
XX
DT 25-MAR-2003 (updated)
DT 21-APR-1994 (first entry)
XX
DE Staphylococcus enterotoxin B mutant BR-210.
XX
KW Staphylococcus; enterotoxin B; SEB; super antigen; SAG; receptor;
KW random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
KW autoimmune disease.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FH Region 97..157
FT /note= "Region undisclosed in the specification"
FT Region 162..186
FT /note= "Region undisclosed in the specification"
XX
PN WO9314634-A1.
XX
PD 05-AUG-1993.
XX
PF 28-JAN-1993; 93WO-US00839.
XX
PR 28-JAN-1992; 92US-0827540.
XX
PA (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
PI Kappler JW, Marrack P;
XX
DR WPI; 1993-320314/40.
XX
PT Preventing or treating toxic effects of super antigens - by admin. of
PT new modified or mutated super antigen which induces antibodies but
PT not T-cell activation
XX
PS Disclosure; Fig 3; 54pp; English.
XX
CC The sequences given in AAR41990-2013 are mutant versions of the
CC Staphylococcus enterotoxin B (SEB) super antigen (SAG). These mutants

CC generated by random mutation by PCR, may be used to modify the T-cell
CC response elicited by an antigen, by interacting with specific Vbeta
CC elements of the T-cell receptors. This T-cell response modulation
CC may be useful in a wide variety of autoimmune diseases, where self-
CC reactive T-cells may be activated by SAG that bind to particular
CC Vbeta types.
CC (Updated on 25-MAR-2003 to correct PN field.)
CC
XX
SQ Sequence 190 AA;
XX
Query Match 1.3%; Score 17; DB 14; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 NNNNNNNNNNNNNNNNNI 758
DB 171 NNNNNNNNNNNNNNNNNI 187
XX
RESULT 4
AAR41992
ID AAR41992 standard; protein; 190 AA.
XX
AC AAR41992;
XX
DT 25-MAR-2003 (updated)
DT 21-APR-1994 (first entry)
XX
DE Staphylococcus enterotoxin B mutant BR-257.
XX
KW Staphylococcus; enterotoxin B; SEB; super antigen; SAG; receptor;
KW random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
KW autoimmune disease.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FH Region 97..157
FT /note= "Region undisclosed in the specification"
FT Region 162..186
FT /note= "Region undisclosed in the specification"
XX
PN WO9314634-A1.
XX
PD 05-AUG-1993.
XX
PF 28-JAN-1993; 93WO-US00839.
XX
PR 28-JAN-1992; 92US-0827540.
XX
PA (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
PI Kappler JW, Marrack P;
XX
DR WPI; 1993-320314/40.
XX
PT Preventing or treating toxic effects of super antigens - by admin. of
PT new modified or mutated super antigen which induces antibodies but
PT not T-cell activation
XX
PS Disclosure; Fig 3; 54pp; English.
XX
CC The sequences given in AAR41990-2013 are mutant versions of the
CC Staphylococcus enterotoxin B (SEB) super antigen (SAG). These mutants
CC generated by random mutation by PCR, may be used to modify the T-cell
CC response elicited by an antigen, by interacting with specific Vbeta
CC elements of the T-cell receptors. This T-cell response modulation
CC may be useful in a wide variety of autoimmune diseases, where self-
CC reactive T-cells may be activated by SAG that bind to particular
CC Vbeta types.
CC (Updated on 25-MAR-2003 to correct PN field.)
CC
XX
SQ Sequence 190 AA;
XX

Query Match 1.3%; Score 17; DB 14; Length 190;
 Best Local Similarity 100.0%; Pred. No. 1.2e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNNI 758
 |||||
 DB 171 NNNNNNNNNNNNNNNNNI 187

RESULT 5

AAR41993
 ID AAR41993 standard; protein; 190 AA.

XX AAR41993;

XX AC 25-MAR-2003 (updated)
 DT 21-APR-1994 (first entry)

XX DE Staphylococcus enterotoxin B mutant BR-291.

XX KM Staphylococcus; enterotoxin B; SEB; super antigen; SAg; receptor;
 KW random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
 XX autoimmune disease.

XX OS Staphylococcus aureus.

XX FH Key Location/Qualifiers
 FT Region 97..157
 FT /note= "Region undisclosed in the specification"

FT Region 162..186

FT /note= "Region undisclosed in the specification"

XX W09314634-A1.

XX PN 05-AUG-1993.

XX PD 28-JAN-1993; 93WO-US00839.

XX PF 28-JAN-1992; 92US-0827540.

XX PR (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.

XX PA Kappler JM, Marrack P;

XX PI WPI; 1993-320314/40.

XX DR Preventing or treating toxic effects of super antigens - by admin. of
 XX PT new modified or mutated super antigen which induces antibodies but
 XX PT not T-cell activation

XX PS Disclosure; Fig 3; 54pp; English.

XX CC The sequences given in AAR41990-2013 are mutant versions of the

XX CC Staphylococcus enterotoxin B (SEB) super antigen (SAg). These mutants

XX CC generated by random mutation by PCR, may be used to modify the T-cell

XX CC response elicited by an antigen, by interacting with specific Vbeta

XX CC elements of the T-cell receptors. This T-cell response modulation

XX CC may be useful in a wide variety of autoimmune diseases, where self-

XX CC reactive T-cells may be activated by SAg that bind to particular

XX CC Vbeta types.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 190 AA;

QY 742 NNNNNNNNNNNNNNNNNI 758
 |||||

DB 171 NNNNNNNNNNNNNNNNNI 187

RESULT 6
 AAR41994
 ID AAR41994 standard; protein; 190 AA.

XX AAR41994;

XX AC 25-MAR-2003 (updated)
 DT 21-APR-1994 (first entry)

XX DE Staphylococcus enterotoxin B mutant BC-6.

XX KM Staphylococcus; enterotoxin B; SEB; super antigen; SAg; receptor;
 KW random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
 XX autoimmune disease.

XX OS Staphylococcus aureus.

XX FH Key Location/Qualifiers
 FT Region 97..157
 FT /note= "Region undisclosed in the specification"

FT Region 162..186

FT /note= "Region undisclosed in the specification"

XX W09314634-A1.

XX PN 05-AUG-1993.

XX PD 28-JAN-1993; 93WO-US00839.

XX PF 28-JAN-1992; 92US-0827540.

XX PR (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.

XX PA Kappler JM, Marrack P;

XX PI WPI; 1993-320314/40.

XX DR Preventing or treating toxic effects of super antigens - by admin. of
 XX PT new modified or mutated super antigen which induces antibodies but
 XX PT not T-cell activation

XX PS Disclosure; Fig 3; 54pp; English.

XX CC The sequences given in AAR41990-2013 are mutant versions of the

XX CC Staphylococcus enterotoxin B (SEB) super antigen (SAg). These mutants

XX CC generated by random mutation by PCR, may be used to modify the T-cell

XX CC response elicited by an antigen, by interacting with specific Vbeta

XX CC elements of the T-cell receptors. This T-cell response modulation

XX CC may be useful in a wide variety of autoimmune diseases, where self-

XX CC reactive T-cells may be activated by SAg that bind to particular

XX CC Vbeta types.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 190 AA;

QY 742 NNNNNNNNNNNNNNNNNI 758
 |||||

DB 171 NNNNNNNNNNNNNNNNNI 187

RESULT 7
 AAR41995
 ID AAR41995 standard; protein; 190 AA.

XX AAR41995;

XX AC 25-MAR-2003 (updated)
 DT 21-APR-1994 (first entry)


```

XX Staphylococcus enterotoxin B mutant BC-66.
DE
XX
XX Staphylococcus; enterotoxin B; SEB; super antigen; SAg; receptor;
KW random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
KW autoimmune disease.
XX
XX OS Staphylococcus aureus.
XX
XX FH Key Location/Qualifiers
XX FT Region 97..157
XX FT /note= "Region undisclosed in the specification"
FT 162..186
FT /note= "Region undisclosed in the specification"
XX
XX PN WO9314634-A1.
XX
XX PD 05-AUG-1993.
XX
XX PF 28-JAN-1993; 93WO-US00839.
PR 28-JAN-1992; 92US-0827540.
XX
XX PA (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
XX PI Kappler JW, Marrack P;
XX
XX DR WPJ; 1993-320314/40.
XX
XX PT Preventing or treating toxic effects of super antigens - by admin. of
PT new modified or mutated super antigen which induces antibodies but
PT not T-cell activation
XX
XX PS Disclosure; Fig 3; 54pp; English.
XX
XX The sequences given in AAR41990-2013 are mutant versions of the
Staphylococcus enterotoxin B (SEB) super antigen (SAg). These mutants
generated by random mutation by PCR, may be used to modify the T-cell
response elicited by an antigen, by interacting with specific Vbeta
elements of the T-cell receptors. This T-cell response modulation
may be useful in a wide variety of autoimmune diseases, where self-
reactive T-cells may be activated by SAg that bind to particular
Vbeta types.
CC
CC (Updated on 25-MAR-2003 to correct PN field.)
CC
SQ Sequence 190 AA;
Query Match: 1.3%; Score 17; DB 14; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 742 NNNNNNNNNNNNNNNNI 758
Db 171 NNNNNNNNNNNNNNNNNI 187
RESULT 8
AAR41996
ID AAR41996 standard; protein; 190 AA.
XX
XX AAR41996;
XX AC
XX XX 25-MAR-2003 (updated)
DT 21-APR-1994 (first entry)
XX DE Staphylococcus enterotoxin B mutant BC-88.
XX
XX KW Staphylococcus; enterotoxin B; SEB; super antigen; SAg; receptor;
KW random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
XX autoimmune disease.
XX
XX SS Staphylococcus aureus.
XX
```

FT	Key	Location/Qualifiers
PH	Region	97..157
FT	Region	/note="Region undisclosed in the specification"
FT	Region	162..186
FT	Region	/note="Region undisclosed in the specification"
XX	PN	W09314634-A1.
XX	PD	05-AUG-1993.
XX	PF	28-JAN-1993; 93WO-US00839.
XX	PR	28-JAN-1992; 92US-0627540.
XX	PA	(NAME-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX	PI	Kappler JW, Marrack P;
XX	DR	WPI, 1993-320314/40.
XX	PT	Preventing or treating toxic effects of super antigens - by admin. of
XX	PT	new modified or mutated super antigen which induces antibodies but
XX	PT	not T-cell activation
XX	PS	Disclosure; Fig 3; 54pp; English.
XX	CC	The sequences given in AAR41990-2013 are mutant versions of the
XX	CC	Staphylococcus enterotoxin B (SEB) super antigen (SAG). These mutants
XX	CC	generated by random mutation by PCR, may be used to modify the T-cell
XX	CC	response elicited by an antigen, by interacting with specific Vbeta
XX	CC	elements of the T-cell receptors. This T-cell response modulation
XX	CC	may be useful in a wide variety of autoimmune diseases, where self-
XX	CC	reactive T-cells may be activated by SAg that bind to particular
XX	CC	Vbeta types.
XX	CC	(Updated on 25-MAR-2003 to correct PN field.)
SO	Sequence	190 AA;
QY	Query Match	1.3%; Score 17; DB 14; Length 190;
DB	Best Local Similarity	100.0%; Pred.No.1.2e-07;
DB	Matches 17; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
DB	742 NNNNNNNNNNNNNNNNT	758
DB	171 NNNNNNNNNNNNNNNNNI	187
RESULT 9		
AAR41997		
ID	AAR41997 standard; protein;	190 AA.
AC	AAR41997;	
XX		
DT	25-MAR-2003 (updated)	
DT	21-APR-1994 (first entry)	
XX		
DE	Staphylococcus enterotoxin B mutant BR-358.	
XX		
KW	Staphylococcus; enterotoxin B; SEB; super antigen; SAg; receptor;	
KW	random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;	
KW	autoimmune disease.	
XX		
OS	Staphylococcus aureus.	
XX		
PH	Key	Location/Qualifiers
FT	Region	97..157
FT	Region	/note="Region undisclosed in the specification"
FT	Region	162..186
FT	Region	/note="Region undisclosed in the specification"
XX	PN	W09314634-A1.
XX	PD	05-AUG-1993.

XX 28-JAN-1993; 93WO-US00839.
 XX
 XX 28-JAN-1992; 92US-0827540.
 XX
 XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
 XX
 XX Kappler JW, Marrack P;
 XX
 XX WPI; 1993-320314/40.
 XX
 XX Preventing or treating toxic effects of super antigens - by admin. of
 XX new modified or mutated super antigen which induces antibodies but
 XX not T-cell activation
 XX
 XX
 XX Disclosure; Fig 3; 54pp; English.
 XX
 XX The sequences given in AAR41990-2013 are mutant versions of the
 XX Staphylococcus enterotoxin B (SEB) super antigen (Sag). These mutants
 XX generated by random mutation by PCR, may be used to modify the T-cell
 XX response elicited by an antigen, by interacting with specific Vbeta
 XX elements of the T-cell receptors. This T-cell response modulation
 XX may be useful in a wide variety of autoimmune diseases, where self-
 XX reactive T-cells may be activated by Sag that bind to particular
 XX Vbeta types.
 XX (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 190 AA;
 XX

Query Match 1.3%; Score 17; DB 14; Length 190;
 Best Local Similarity 100.0%; Pred. No. 1.2e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 742 NNNNNNNNNNNNNNNNNI 758
 ||||||||||||||||
 Db 171 NNNNNNNNNNNNNNNNNI 187

RESULT 10
 AAR41998
 ID AAR41998 standard; protein; 190 AA.
 XX
 XX AAR41998;
 XX
 XX 25-MAR-2003 (updated)
 DT 21-APR-1994 (first entry)
 XX
 XX Staphylococcus enterotoxin B mutant BR-374.
 XX
 XX Staphylococcus; enterotoxin B; SEB; super antigen; Sag; receptor;
 KW random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
 KW autoimmune disease.
 XX
 XX Staphylococcus aureus.
 OS
 XX
 XX Key Location/Qualifiers
 FH 97..157
 FT /note= "Region undisclosed in the specification"
 FT 162..186
 FT Region /note= "Region undisclosed in the specification"
 XX
 XX WO9314634-A1.
 PN
 XX
 XX 05-AUG-1993.
 PD
 XX
 XX 28-JAN-1993; 93WO-US00839.
 PF
 XX
 XX 28-JAN-1992; 92US-0827540.
 PR
 XX
 XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
 PA
 XX Kappler JW, Marrack P;
 PI
 XX

DR WPI; 1993-320314/40.
 XX
 XX Preventing or treating toxic effects of super antigens - by admin. of
 XX new modified or mutated super antigen which induces antibodies but
 XX not T-cell activation
 XX
 XX
 XX Disclosure; Fig 3; 54pp; English.
 XX
 XX The sequences given in AAR41990-2013 are mutant versions of the
 XX Staphylococcus enterotoxin B (SEB) super antigen (Sag). These mutants
 XX generated by random mutation by PCR, may be used to modify the T-cell
 XX response elicited by an antigen, by interacting with specific Vbeta
 XX elements of the T-cell receptors. This T-cell response modulation
 XX may be useful in a wide variety of autoimmune diseases, where self-
 XX reactive T-cells may be activated by Sag that bind to particular
 XX Vbeta types.
 XX (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 190 AA;
 XX

Query Match 1.3%; Score 17; DB 14; Length 190;
 Best Local Similarity 100.0%; Pred. No. 1.2e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 742 NNNNNNNNNNNNNNNNNI 758
 ||||||||||||||||
 Db 171 NNNNNNNNNNNNNNNNNI 187

RESULT 11
 AAR41999
 ID AAR41999 standard; protein; 190 AA.
 XX
 XX AAR41999;
 XX
 XX 25-MAR-2003 (updated)
 DT 21-APR-1994 (first entry)
 XX
 XX Staphylococcus enterotoxin B mutant BA-3.
 XX
 XX Staphylococcus; enterotoxin B; SEB; super antigen; Sag; receptor;
 KW random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
 KW autoimmune disease.
 XX
 XX Staphylococcus aureus.
 OS
 XX
 XX Key Location/Qualifiers
 FH 97..157
 FT /note= "Region undisclosed in the specification"
 FT 162..186
 FT Region /note= "Region undisclosed in the specification"
 XX
 XX WO9314634-A1.
 PN
 XX
 XX 05-AUG-1993.
 PD
 XX
 XX 28-JAN-1993; 93WO-US00839.
 PF
 XX
 XX 28-JAN-1992; 92US-0827540.
 PR
 XX
 XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
 PA
 XX Kappler JW, Marrack P;
 PI
 XX
 XX WPI; 1993-320314/40.
 DR
 XX
 XX Preventing or treating toxic effects of super antigens - by admin. of
 XX new modified or mutated super antigen which induces antibodies but
 XX not T-cell activation
 XX
 XX
 XX Disclosure; Fig 3; 54pp; English.
 XX
 XX The sequences given in AAR41990-2013 are mutant versions of the


```
CC Staphylococcus enterotoxin B (SEB) super antigen (Sag). These mutants
CC generated by random mutation by PCR, may be used to modify the T-cell
CC response elicited by an antigen, by interacting with specific Vbeta
CC elements of the T-cell receptors. This T-cell response modulation
CC may be useful in a wide variety of autoimmune diseases, where self-
CC reactive T-cells may be activated by Sag that bind to particular
CC Vbeta types.
CC (Updated on 25-MAR-2003 to correct PN field.)
SQ Sequence 190 AA;

Query Match 1.3%; Score 17; DB 14; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNNI 758
DB 171 NNNNNNNNNNNNNNNNNI 187

RESULT 12
AAR42000
ID AAR42000 standard; protein; 190 AA.
AC AAR42000;
XX
XX
XX 25-MAR-2003 (updated)
DT 21-APR-1994 (first entry)
XX
XX Staphylococcus enterotoxin B mutant BA-15.
DE
XX Staphylococcus; enterotoxin B; SEB; super antigen; Sag; receptor;
KM random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
KM autoimmune disease.
XX
XX Staphylococcus aureus.
OS
XX
XX Key Location/Qualifiers
FH Region 97..157
FT /note= "Region undisclosed in the specification"
FT Region 162..186
FT /note= "Region undisclosed in the specification"
XX
XX WO9314634-A1.
XX
XX
XX 05-AUG-1993.
PD
XX
XX 28-JAN-1993; 93WO-US00839.
PF
XX
XX 28-JAN-1992; 92US-0827540.
PR
XX
XX (NAME-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
PA
XX Kappler JW, Marrack P;
PI
XX
XX WPI; 1993-320314/40.
XX
XX Preventing or treating toxic effects of super antigens - by admin. of
PT new modified or mutated super antigen which induces antibodies but
PT not T-cell activation
XX
XX Disclosure; Fig 3; 54pp; English.
XX
XX The sequences given in AAR41990-2013 are mutant versions of the
CC Staphylococcus enterotoxin B (SEB) super antigen (Sag). These mutants
CC generated by random mutation by PCR, may be used to modify the T-cell
CC response elicited by an antigen, by interacting with specific Vbeta
CC elements of the T-cell receptors. This T-cell response modulation
CC may be useful in a wide variety of autoimmune diseases, where self-
CC reactive T-cells may be activated by Sag that bind to particular
CC Vbeta types.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
```

```
SQ Sequence 190 AA;

Query Match 1.3%; Score 17; DB 14; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNNI 758
DB 171 NNNNNNNNNNNNNNNNNI 187

RESULT 13
AAR42001
ID AAR42001 standard; protein; 190 AA.
AC AAR42001;
XX
XX
XX 25-MAR-2003 (updated)
DT 21-APR-1994 (first entry)
XX
XX Staphylococcus enterotoxin B mutant BA-24.
DE
XX Staphylococcus; enterotoxin B; SEB; super antigen; Sag; receptor;
KM random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
KM autoimmune disease.
XX
XX Staphylococcus aureus.
OS
XX
XX Key Location/Qualifiers
FH Region 97..157
FT /note= "Region undisclosed in the specification"
FT Region 162..186
FT /note= "Region undisclosed in the specification"
XX
XX WO9314634-A1.
XX
XX
XX 05-AUG-1993.
PD
XX
XX 28-JAN-1993; 93WO-US00839.
PF
XX
XX 28-JAN-1992; 92US-0827540.
PR
XX
XX (NAME-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
PA
XX Kappler JW, Marrack P;
PI
XX
XX WPI; 1993-320314/40.
XX
XX Preventing or treating toxic effects of super antigens - by admin. of
PT new modified or mutated super antigen which induces antibodies but
PT not T-cell activation
XX
XX Disclosure; Fig 3; 54pp; English.
XX
XX The sequences given in AAR41990-2013 are mutant versions of the
CC Staphylococcus enterotoxin B (SEB) super antigen (Sag). These mutants
CC generated by random mutation by PCR, may be used to modify the T-cell
CC response elicited by an antigen, by interacting with specific Vbeta
CC elements of the T-cell receptors. This T-cell response modulation
CC may be useful in a wide variety of autoimmune diseases, where self-
CC reactive T-cells may be activated by Sag that bind to particular
CC Vbeta types.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 190 AA;

Query Match 1.3%; Score 17; DB 14; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNNI 758
DB 171 NNNNNNNNNNNNNNNNNI 187
```


RESULT 14
AAR42002
ID AAR42002 standard; protein, 190 AA.
XX
AC AAR42002;
XX
DT 25-MAR-2003 (updated)
DT 21-APR-1994 (first entry)
XX
DE Staphylococcus enterotoxin B mutant BA-31.
XX
KW Staphylococcus; enterotoxin B; SEB; super antigen; SAg; receptor; random mutation; PCR; modify; T-cell; response; antigen; Vbeta element; autoimmune disease.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Region 97..157
FT /note= "Region undisclosed in the specification"
FT Region 162..186
FT /note= "Region undisclosed in the specification"
XX
PN WO9314634-A1.
XX
PD 05-AUG-1993.
XX
PE 28-JAN-1993; 93WO-US00839.
XX
PR 28-JAN-1992; 92US-0827540.
XX
PA (NAME-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
PI Kappler JW, Marrack P;
DR WPT, 1993-320314/40.
XX
PT Preventing or treating toxic effects of super antigens - by admin. of new modified or mutated super antigen which induces antibodies but not T-cell activation
XX
PS Disclosure; Fig 3; 5app; English.
XX
CC The sequences given in AAR41990-2013 are mutant versions of the Staphylococcus enterotoxin B (SEB) super antigen (SAg). These mutants generated by random mutation by PCR, may be used to modify the T-cell response elicited by an antigen, by interacting with specific Vbeta elements of the T-cell receptors. This T-cell response modulation may be useful in a wide variety of autoimmune diseases, where self- reactive T-cells may be activated by SAg that bind to particular Vbeta types.
CC
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 190 AA:
Query Match 1.3%; Score 17; DB 14; Length 190; Best Local Similarity 100.0%; Pred. No. 1.2e-07; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
OY 742 NNNNNNNNNNNNNNNNI 758 Db 171 NNNNNNNNNNNNNNNNI 187
RESULT 15
AAR42003
ID AAR42003 standard; protein, 190 AA.
XX
AC AAR42003;
XX
DT 25-MAR-2003 (updated)

DT	21-APR-1994	(first entry)
DE	Staphylococcus enterotoxin B mutant BA-50.	
KW	Staphylococcus; enterotoxin B; SEB; super antigen; SAg; receptor;	
KM	random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;	
KX	autoimmune disease.	
OS	Staphylococcus aureus.	
FH	Key	
FT	Region	
FT	/note= "Region undisclosed in the specification"	
FT	Region	
PN	/note= "Region undisclosed in the specification"	
PN	WO9314634-A1.	
XX		
PD	05-AUG-1993.	
PF	28-JAN-1993;	
PR	28-JAN-1992;	
PA	(NAME-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.	
PI	Kappler JW, Marrack P;	
DR	WPI, 1993-320314/40.	
PT	Preventing or treating toxic effects of super antigens - by admin. of	
PT	new modified or mutated super antigen which induces antibodies but	
PT	not T-cell activation	
PS	Disclosure; Fig 3; 54pp; English.	
CC	The sequences given in AAR41990-2013 are mutant versions of the	
CC	Staphylococcus enterotoxin B (SEB) super antigen (SAg). These mutants	
CC	generated by random mutation by PCR, may be used to modify the T-cell	
CC	response elicited by an antigen, by interacting with specific Vbeta	
CC	elements of the T-cell receptors. This T-cell response modulation	
CC	may be useful in a wide variety of autoimmune diseases, where self-	
CC	reactive T-cells may be activated by SAg that bind to particular	
CC	Vbeta types.	
CC	(updated on 25-MAR-2003 to correct PN field.)	
SQ	Sequence 190 AA;	
Query Match	1.3%; Score 17; DB 14; Length 190;	
Best Local Similarity	100.0%; Pred. No. 1.2e-07;	
Matches 17; Conservative	0; Mismatches 0; Gaps 0;	
OY	742 NNNNNNNNNNNNNNNNI 758 	
Db	171 NNNNNNNNNNNNNNNNI 187	
RESULT 16		
AAR42004	ID AAR42004 standard; protein; 190 AA.	
AC	AAR42004;	
DT	25-MAR-2003 (updated)	
DT	21-APR-1994 (first entry)	
DE	Staphylococcus enterotoxin B mutant BA-53.	
KW	Staphylococcus; enterotoxin B; SEB; super antigen; SAg; receptor;	
KM	random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;	
KX	autoimmune disease.	
OS	Staphylococcus aureus.	


```

XX Key Location/Qualifiers
FH Region 97..157
FT /note="Region undisclosed in the specification"
FT 162..186
FT /note="Region undisclosed in the specification"
XX
XX MO9314634-A1.
XX
XX PD 05-AUG-1993.
XX
XX PF 28-JAN-1993; 93WO-US00839.
XX
XX PR 28-JAN-1992; 92US-0827540.
XX
XX PA (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
XX PI Kappler JW, Marrack P;
XX
XX DR WPI; 1993-320314/40.
XX
XX PT Preventing or treating toxic effects of super antigens - by admin. of
XX new modified or mutated super antigen which induces antibodies but
XX not T-cell activation
XX
XX PS Disclosure; Fig 3; 54pp; English.
XX
XX CC The sequences given in AAR41990-2013 are mutant versions of the
XX Staphylococcus enterotoxin B (SEB) super antigen (Sag). These mutants
XX generated by random mutation by PCR, may be used to modify the T-cell
XX response elicited by an antigen, by interacting with specific Vbeta
XX elements of the T-cell receptors. This T-cell response modulation
XX may be useful in a wide variety of autoimmune diseases, where self-
XX reactive T-cells may be activated by Sag that bind to particular
XX Vbeta types.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 190 AA;
XX
XX
XX Query Match 1.3%; Score 17; DB 14; Length 190;
XX Best Local Similarity 100.0%; Pred.No.1.2e-07;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 742 NNNNNNNNNNNNNNNNNI 758
XX |||||
XX 171 NNNNNNNNNNNNNNNNNI 187
XX
XX
XX RESULT 17
XX AAR42005
XX ID AAR42005 standard; protein; 190 AA.
XX
XX AC AAR42005;
XX
XX DT 25-MAR-2003 (updated)
XX DT 21-APR-1994 (first entry)
XX
XX DE Staphylococcus enterotoxin B mutant BA-62.
XX
XX KW Staphylococcus; enterotoxin B; SEB; super antigen; Sag; receptor;
XX random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
XX autoimmune disease.
XX
XX OS Staphylococcus aureus.
XX
XX FH Key Location/Qualifiers
XX FT 97..157
XX FT /note="Region undisclosed in the specification"
XX FT 162..186
XX FT /note="Region undisclosed in the specification"
XX
XX PA (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
XX PI Kappler JW, Marrack P;
XX
XX MO9314634-A1.
XX

```

```

PD 05-AUG-1993.
XX
XX XX 28-JAN-1993; 93WO-US00839.
XX
XX XX 28-JAN-1992; 92US-0827540.
XX
XX PA (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
XX PI Kappler JW, Marrack P;
XX
XX DR WPI; 1993-320314/40.
XX
XX PT Preventing or treating toxic effects of super antigens - by admin. of
XX new modified or mutated super antigen which induces antibodies but
XX not T-cell activation
XX
XX PS Disclosure; Fig 3; 54pp; English.
XX
XX CC The sequences given in AAR41990-2013 are mutant versions of the
XX Staphylococcus enterotoxin B (SEB) super antigen (Sag). These mutants
XX generated by random mutation by PCR, may be used to modify the T-cell
XX response elicited by an antigen, by interacting with specific Vbeta
XX elements of the T-cell receptors. This T-cell response modulation
XX may be useful in a wide variety of autoimmune diseases, where self-
XX reactive T-cells may be activated by Sag that bind to particular
XX Vbeta types.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 190 AA;
XX
XX
XX Query Match 1.3%; Score 17; DB 14; Length 190;
XX Best Local Similarity 100.0%; Pred.No.1.2e-07;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 742 NNNNNNNNNNNNNNNNNI 758
XX |||||
XX 171 NNNNNNNNNNNNNNNNNI 187
XX
XX
XX RESULT 18
XX AAR42006
XX ID AAR42006 standard; protein; 190 AA.
XX
XX AC AAR42006;
XX
XX DT 25-MAR-2003 (updated)
XX DT 21-APR-1994 (first entry)
XX
XX DE Staphylococcus enterotoxin B mutant BB-14.
XX
XX KW Staphylococcus; enterotoxin B; SEB; super antigen; Sag; receptor;
XX random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
XX autoimmune disease.
XX
XX OS Staphylococcus aureus.
XX
XX FH Key Location/Qualifiers
XX FT 97..157
XX FT /note="Region undisclosed in the specification"
XX FT 162..186
XX FT /note="Region undisclosed in the specification"
XX
XX PA (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
XX PI Kappler JW, Marrack P;
XX
XX MO9314634-A1.
XX
XX PD 05-AUG-1993.
XX
XX PF 28-JAN-1993; 93WO-US00839.
XX
XX PR 28-JAN-1992; 92US-0827540.
XX
XX PA (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
XX PI Kappler JW, Marrack P;
XX

```



```
XX DR WPI; 1993-320314/40.
XX PT Preventing or treating toxic effects of super antigens - by admin. of
XX PT new modified or mutated super antigen which induces antibodies but
XX PT not T-cell activation
XX PS Disclosure; Fig 3; 54pp; English.
XX CC The sequences given in AAR41990-2013 are mutant versions of the
XX CC Staphylococcus enterotoxin B (SEB) super antigen (Sag). These mutants
XX CC generated by random mutation by PCR, may be used to modify the T-cell
XX CC response elicited by an antigen, by interacting with specific Vbeta
XX CC elements of the T-cell receptors. This T-cell response modulation
XX CC may be useful in a wide variety of autoimmune diseases, where self-
XX CC reactive T-cells may be activated by Sag that bind to particular
XX CC Vbeta types.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 190 AA;
XX QY Query Match 1.3%; Score 17; DB 14; Length 190;
XX DB Best Local Similarity 100.0%; Pred. No. 1.2e-07;
XX DB Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX DB 742 NNNNNNNNNNNNNNNNNI 758
XX DB 171 NNNNNNNNNNNNNNNNNI 187

RESULT 19
AAR42007
ID AAR42007 standard; protein; 190 AA.
XX AC AAR42007;
XX DT 25-MAR-2003 (updated)
XX DT 21-APR-1994 (first entry)
XX DE Staphylococcus enterotoxin B mutant BB-21.
XX KM Staphylococcus enterotoxin B; SEB; super antigen; Sag; receptor;
XX KM random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
XX KM autoimmune disease.
XX OS Staphylococcus aureus.
XX FH Key Location/Qualifiers
XX FH Region 97..157
XX FT /note= "Region undisclosed in the specification"
XX FT Region 162..186
XX FT /note= "Region undisclosed in the specification"
XX PN MO9314634-A1.
XX PD 05-AUG-1993.
XX PF 28-JAN-1993; 93WO-US00839.
XX PR 28-JAN-1992; 92US-0827540.
XX PA (NAME-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX PI Kappler JW, Marrack P;
XX DR WPI; 1993-320314/40.
XX PT Preventing or treating toxic effects of super antigens - by admin. of
XX PT new modified or mutated super antigen which induces antibodies but
XX PT not T-cell activation
XX PS Disclosure; Fig 3; 54pp; English.
XX CC
```

```
CC CC The sequences given in AAR41990-2013 are mutant versions of the
CC CC Staphylococcus enterotoxin B (SEB) super antigen (Sag). These mutants
CC CC generated by random mutation by PCR, may be used to modify the T-cell
CC CC response elicited by an antigen, by interacting with specific Vbeta
CC CC elements of the T-cell receptors. This T-cell response modulation
CC CC may be useful in a wide variety of autoimmune diseases, where self-
CC CC reactive T-cells may be activated by Sag that bind to particular
CC CC Vbeta types.
CC CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 190 AA;
XX QY Query Match 1.3%; Score 17; DB 14; Length 190;
XX DB Best Local Similarity 100.0%; Pred. No. 1.2e-07;
XX DB Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX DB 742 NNNNNNNNNNNNNNNNNI 758
XX DB 171 NNNNNNNNNNNNNNNNNI 187

RESULT 20
AAR42008
ID AAR42008 standard; protein; 190 AA.
XX AC AAR42008;
XX DT 25-MAR-2003 (updated)
XX DT 21-APR-1994 (first entry)
XX DE Staphylococcus enterotoxin B mutant BB-47.
XX KM Staphylococcus enterotoxin B; SEB; super antigen; Sag; receptor;
XX KM random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
XX KM autoimmune disease.
XX OS Staphylococcus aureus.
XX FH Key Location/Qualifiers
XX FH Region 97..157
XX FT /note= "Region undisclosed in the specification"
XX FT Region 162..186
XX FT /note= "Region undisclosed in the specification"
XX PN MO9314634-A1.
XX PD 05-AUG-1993.
XX PF 28-JAN-1993; 93WO-US00839.
XX PR 28-JAN-1992; 92US-0827540.
XX PA (NAME-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX PI Kappler JW, Marrack P;
XX DR WPI; 1993-320314/40.
XX PT Preventing or treating toxic effects of super antigens - by admin. of
XX PT new modified or mutated super antigen which induces antibodies but
XX PT not T-cell activation
XX PS Disclosure; Fig 3; 54pp; English.
XX CC The sequences given in AAR41990-2013 are mutant versions of the
XX CC Staphylococcus enterotoxin B (SEB) super antigen (Sag). These mutants
XX CC generated by random mutation by PCR, may be used to modify the T-cell
XX CC response elicited by an antigen, by interacting with specific Vbeta
XX CC elements of the T-cell receptors. This T-cell response modulation
XX CC may be useful in a wide variety of autoimmune diseases, where self-
XX CC reactive T-cells may be activated by Sag that bind to particular
XX CC Vbeta types.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX CC
```



```
XX SQ Sequence 190 AA;
Query Match 1.3%; Score 17; DB 14; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 NNNNNNNNNNNNNNNNNI 758
DB 171 NNNNNNNNNNNNNNNNNI 187

RESULT 21
AAR42009
ID AAR42009 standard; protein; 190 AA.
XX AAR42009;
XX
XX 25-MAR-2003 (updated)
XX 21-APR-1994 (first entry)
XX
XX Staphylococcus enterotoxin B mutant BR-30.
XX
XX Staphylococcus enterotoxin B; SEB; super antigen; SAg; receptor;
XX random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
XX autoimmune disease.
XX
XX Staphylococcus aureus.
XX
XX Key Location/Qualifiers
XX Region 97..157
XX /note="Region undisclosed in the specification"
XX FT 162..186
XX /note="Region undisclosed in the specification"
XX
XX MO9314634-A1.
XX
XX 05-AUG-1993.
XX
XX 28-JAN-1993; 93WO-US00839.
XX
XX 28-JAN-1992; 92US-0827540.
XX
XX (NAME-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
XX Kaplier JW, Marrack P;
XX
XX WPI; 1993-320314/40.
XX
XX Preventing or treating toxic effects of super antigens - by admin. of
XX new modified or mutated super antigen which induces antibodies but
XX not T-cell activation
XX
XX Disclosure; Fig 3; 54pp; English.
XX
XX The sequences given in AAR41990-2013 are mutant versions of the
XX Staphylococcus enterotoxin B (SEB) super antigen (SAg). These mutants
XX generated by random mutation by PCR, may be used to modify the T-cell
XX response elicited by an antigen, by interacting with specific Vbeta
XX elements of the T-cell receptors. This T-cell response modulation
XX may be useful in a wide variety of autoimmune diseases, where self-
XX reactive T-cells may be activated by SAg that bind to particular
XX Vbeta types.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 190 AA;
Query Match 1.3%; Score 17; DB 14; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 NNNNNNNNNNNNNNNNNI 758
DB 171 NNNNNNNNNNNNNNNNNI 187
```

```
DB 171 NNNNNNNNNNNNNNNNNI 187

RESULT 22
AAR42010
ID AAR42010 standard; protein; 190 AA.
XX AAR42010;
XX
XX 25-MAR-2003 (updated)
XX 21-APR-1994 (first entry)
XX
XX Staphylococcus enterotoxin B mutant BR-311.
XX
XX Staphylococcus enterotoxin B; SEB; super antigen; SAg; receptor;
XX random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
XX autoimmune disease.
XX
XX Staphylococcus aureus.
XX
XX Key Location/Qualifiers
XX Region 97..157
XX /note="Region undisclosed in the specification"
XX FT 162..186
XX /note="Region undisclosed in the specification"
XX
XX MO9314634-A1.
XX
XX 05-AUG-1993.
XX
XX 28-JAN-1993; 93WO-US00839.
XX
XX 28-JAN-1992; 92US-0827540.
XX
XX (NAME-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
XX Kaplier JW, Marrack P;
XX
XX WPI; 1993-320314/40.
XX
XX Preventing or treating toxic effects of super antigens - by admin. of
XX new modified or mutated super antigen which induces antibodies but
XX not T-cell activation
XX
XX Disclosure; Fig 3; 54pp; English.
XX
XX The sequences given in AAR41990-2013 are mutant versions of the
XX Staphylococcus enterotoxin B (SEB) super antigen (SAg). These mutants
XX generated by random mutation by PCR, may be used to modify the T-cell
XX response elicited by an antigen, by interacting with specific Vbeta
XX elements of the T-cell receptors. This T-cell response modulation
XX may be useful in a wide variety of autoimmune diseases, where self-
XX reactive T-cells may be activated by SAg that bind to particular
XX Vbeta types.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 190 AA;
Query Match 1.3%; Score 17; DB 14; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 NNNNNNNNNNNNNNNNNI 758
DB 171 NNNNNNNNNNNNNNNNNI 187

RESULT 23
AAR42011
ID AAR42011 standard; protein; 190 AA.
XX AAR42011;
XX
```


DT	25-MAR-2003	(updated)
DT	21-APR-1994	(first entry)
XX		
XX	Staphylococcus enterotoxin B mutant BR-474.	
XX		
KW	Staphylococcus; enterotoxin B; SEB; super antigen; SAg; receptor;	
KW	random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;	
KW	autoimmune disease.	
OS		
XX	Staphylococcus aureus.	
XX		
FT	Key	Location/Qualifiers
FT	Region	97..157
FT	Region	/note= "Region undisclosed in the specification"
FT		162..186
PN		/note= "Region undisclosed in the specification"
PD		
XX	WO9314634-A1.	
XX	05-AUG-1993.	
XX		
PF	28-JAN-1993;	93WO-US00839.
PR	28-JAN-1992;	92US-0827540.
XX		
PA	(NAME-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.	
XX		
P1	Kappler JW, Marrack P;	
XX		
DR	WPI, 1993-320314/40.	
PT		
PT	Preventing or treating toxic effects of super antigens - by admin. of	
PT	new modified or mutated super antigen which induces antibodies but	
PT	not T-cell activation	
PS		
PS	Disclosure; Fig 3; 54pp; English.	
XX		
CC	The sequences given in AAR41990-2013 are mutant versions of the	
CC	Staphylococcus enterotoxin B (SEB) super antigen (SAg). These mutants	
CC	generated by random mutation by PCR, may be used to modify the T-cell	
CC	response elicited by an antigen, by interacting with specific Vbeta	
CC	elements of the T-cell receptors. This T-cell response modulation	
CC	may be useful in a wide variety of autoimmune diseases, where self-	
CC	reactive T-cells may be activated by SAgs that bind to particular	
CC	Vbeta types.	
CC	(Updated on 25-MAR-2003 to correct PN field.)	
SQ	Sequence	190 AA;
	Query Match	1.3%; Score 17; DB 14; Length 190;
	Best Local Similarity	100.0%; Pred.No. 1.2e-07;
	Matches 17; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY		
	742 NNNNNNNNNNNNNNNNI	758
DB	171 NNNNNNNNNNNNNNNNI	187
	RESULT 24	
ID	AAR42012 standard; protein; 190 AA.	
XX		
AC	AAR42012;	
XX		
DT	25-MAR-2003 (updated)	
DT	21-APR-1994 (first entry)	
XX		
DE	Staphylococcus enterotoxin B mutant BA-72.	
XX		
KW	Staphylococcus; enterotoxin B; SEB; super antigen; SAg; receptor;	
KW	random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;	
KW	autoimmune disease.	
XX		

OS	Staphylococcus aureus.
XX	
FH	Key Location/Qualifiers
FT	Region 97..157
FT	/note= "Region undisclosed in the specification"
FT	Region 162..186
FT	/note= "Region undisclosed in the specification"
XX	
PN	W09314634-A1.
XX	
PD	05-AUG-1993.
XX	
PF	28-JAN-1993; 93WO-US00839.
XX	
PR	28-JAN-1992; 92US-0827540.
XX	
PA	(NAME-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX	
PI	Kappler JW, Marrack P;
DR	WPI; 1993-320314/40.
XX	
PT	Preventing or treating toxic effects of super antigens - by admin. of
PT	new modified or mutated super antigen which induces antibodies but
PT	not T-cell activation
XX	
P5	Disclosure; Fig 3; 54pp; English.
CC	
CC	The sequences given in AAR41990-2013 are mutant versions of the
CC	Staphylococcus enterotoxin B (SEB) super antigen (SAg). These mutants
CC	generated by random mutation by PCR, may be used to modify the T-cell
CC	response elicited by an antigen, by interacting with specific Vbeta
CC	elements of the T-cell receptors. This T-cell response modulation
CC	may be useful in a wide variety of autoimmune diseases, where self-
CC	reactive T-cells may be activated by SAg that bind to particular
CC	Vbeta types.
CC	(Updated on 25-MAR-2003 to correct PN field.)
SQ	Sequence 190 AA:
	Query Match 1.3%; Score 17; DB 14; Length 190;
	Best Local Similarity 100.0%; Pred. No. 1.2e-07;
	Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	742 NNNNNNNNNNNNNNNNI 758
DB	171 NNNNNNNNNNNNNNNNI 187
RESULT 25	
AAR42013	
ID	AAR42013 standard; protein; 190 AA.
XX	
AC	AAR42013;
DT	25-MAR-2003 (updated)
DT	21-APR-1994 (first entry)
DE	Staphylococcus enterotoxin B mutant BA-267.
XX	
KW	Staphylococcus; enterotoxin B; SEB; super antigen; SAg; receptor;
KW	random mutation; PCR; modify; T-cell; response; antigen; Vbeta element;
XX	autoimmune disease.
OS	Staphylococcus aureus.
XX	
FH	Key Location/Qualifiers
FT	Region 97..157
FT	/note= "Region undisclosed in the specification"
FT	Region 162..186
FT	/note= "Region undisclosed in the specification"
XX	
PN	W09314634-A1.

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155133.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.2%; Score 16; DB 21; Length 141;
Best Local Similarity 100.0%; Pred.No. 8; 6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 741 TNNNNNNNNNNNNNNNN 756
Db 2 TNNNNNNNNNNNNNNNN 17

RESULT 27
AAG49726
ID AAG49726 standard; Protein; 150 AA.
XX
AC AAG49726;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62936.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142290.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.

PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.2%; Score 16; DB 21; Length 150;
Best Local Similarity 100.0%; Pred. No. 9.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 741 TNNNNNNNNNNNNNNNN 756
DB 11 TNNNNNNNNNNNNNNNN 26

RESULT 28

ABP73877
ID ABP73877 standard; Protein; 213 AA.

XX ABP73877;
XX

DT 30-JAN-2003 (first entry)

DE Candida albicans essential protein SEQ ID NO 7714.

XX Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;

KM signal transduction; DNA replication; cell growth;

KM proliferation; Candida albicans; fungicide; antifungal.

XX Candida albicans.

XX WO200253728-A2.

PD 11-JUL-2002.

XX 26-DEC-2001; 2001WO-US49486.

XX 29-DEC-2000; 2000US-259128P.

PR 20-FEB-2001; 2001US-0792024.

PR 22-AUG-2001; 2001US-314050P.

XX (ELIT-) ELITRA PHARM INC.

XX Roemer T, Jjiang B, Boone C, Bussey H, Ohlsen KL;

PI WPI: 2002-566694/60.

DR N-PSDB; ABZ32427.

XX Constructing strains for identifying gene products as effective targets

PT for therapeutic intervention, by inactivating in the strain one allele

PT of a gene and placing other allele of the gene under conditional

XX expression

XX Claim 44; SEQ ID NO 7714; 167pp + Sequence Listing; English.

CC cells in which both alleles of a gene are modified. The diploid fungal
CC cells having both alleles modified are useful for identifying a gene that
CC is essential to the survival or growth of a fungus, a gene that
CC contributes to the virulence and/or pathogenicity of a fungus, a gene
CC that contributes to the resistance of a diploid fungus to an antifungal
CC agent, an antifungal agent that inhibits the growth of a diploid fungus
CC and for identifying a therapeutic agent for treatment of a mammalian
CC disease. (M1) is useful for identifying a compound which modulates the
CC activity of a gene product, preferably enzymatic activity, carbon
CC compound catabolism, biosynthetic, transporter, transcriptional,
CC translational, signal transduction, DNA replication and cell division
CC activity. The method is useful for identifying a compound having the
CC ability to inhibit growth or proliferation of C. albicans cells and for
CC treating infection by C. albicans. The present sequence is that of an
CC essential Candida albicans protein used in the method of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.

XX SQ Sequence 213 AA;

Query Match 1.2%; Score 16; DB 23; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 741 TNNNNNNNNNNNNNNNN 756
DB 172 TNNNNNNNNNNNNNNNN 187

RESULT 29

AAR24796
ID AAR24796 standard; Protein; 256 AA.

XX AAR24796;
XX

DT 25-MAR-2003 (updated)

DT 02-JAN-1992 (first entry)

XX Sequence of fibronectin-derived peptide which binds GPIIb-IIIa.

XX Fibronectin; cell attachment; cell adhesion; adhesive glycoprotein.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "bonded to NH2 gp.or N-terminal SQ of AAs"

XX WO9209200-A1.

XX 11-JUN-1992.

XX 03-DEC-1991; 91WO-US09029.

XX 03-DEC-1990; 90US-0620668.

PR 03-JUL-1991; 91US-0725600.

PR 27-NOV-1991; 91US-0803623.

XX (SCRI) SCRIPPS RES INST.

XX Bowditch R, Ginsberg MH, Plow EF;

XX WPI: 1992-216714/26.

XX New polypeptide(s) derived from human fibronectin - promote cell

PT attachment to substrates, inhibit platelet aggregation and

XX thrombus formation and modulate coagulation and inflammatory
XX response
XX Claim 5; Page 97; 112pp; English.
XX AAQ25592 comprises bps 901-1506 of a human fibronectin (Fn) encoding DNA

PR	03-AUG-1999	99US-0147038
PR	04-AUG-1999	99US-0147204
PR	04-AUG-1999	99US-0147204
PR	04-AUG-1999	99US-0147204
PR	05-AUG-1999	99US-0147302
PR	05-AUG-1999	99US-0147302
PR	06-AUG-1999	99US-0147260
PR	06-AUG-1999	99US-0147303
PR	06-AUG-1999	99US-0147303
PR	06-AUG-1999	99US-0147303
PR	06-AUG-1999	99US-0147303
PR	09-AUG-1999	99US-0147493
PR	09-AUG-1999	99US-0147493
PR	10-AUG-1999	99US-0147935
PR	10-AUG-1999	99US-0148171
PR	11-AUG-1999	99US-0148319
PR	12-AUG-1999	99US-0148341
PR	13-AUG-1999	99US-0148565
PR	13-AUG-1999	99US-0148665
PR	16-AUG-1999	99US-0148664
PR	16-AUG-1999	99US-0149368
PR	17-AUG-1999	99US-0149415
PR	18-AUG-1999	99US-0149416
PR	20-AUG-1999	99US-0149722
PR	20-AUG-1999	99US-0149723
PR	20-AUG-1999	99US-0149929
PR	23-AUG-1999	99US-0149902
PR	23-AUG-1999	99US-0149902
PR	23-AUG-1999	99US-0149902
PR	25-AUG-1999	99US-0150566
PR	25-AUG-1999	99US-0150566
PR	27-AUG-1999	99US-0150884
PR	27-AUG-1999	99US-0151065
PR	27-AUG-1999	99US-0151066
PR	27-AUG-1999	99US-0151080
PR	30-AUG-1999	99US-0151303
PR	30-AUG-1999	99US-0151303
PR	31-SEP-1999	99US-0151438
PR	01-SEP-1999	99US-0151593
PR	07-SEP-1999	99US-0152360
PR	10-SEP-1999	99US-0153070
PR	13-SEP-1999	99US-0153070
PR	15-SEP-1999	99US-0153758
PR	15-SEP-1999	99US-0154018
PR	16-SEP-1999	99US-0154039
PR	20-SEP-1999	99US-0154779
PR	22-SEP-1999	99US-0155119
PR	23-SEP-1999	99US-0155486
PR	24-SEP-1999	99US-0155659
PR	28-SEP-1999	99US-0156458
PR	28-SEP-1999	99US-0156566
PR	04-OCT-1999	99US-0157171
PR	05-OCT-1999	99US-0157753
PR	06-OCT-1999	99US-0157865
PR	07-OCT-1999	99US-0158029
PR	08-OCT-1999	99US-0158232
PR	12-OCT-1999	99US-0158369
PR	13-OCT-1999	99US-0158924
PR	13-OCT-1999	99US-0159294
PR	14-OCT-1999	99US-0159295
PR	14-OCT-1999	99US-0159329
PR	14-OCT-1999	99US-0159330
PR	14-OCT-1999	99US-0159637
PR	18-OCT-1999	99US-0159638
PR	21-OCT-1999	99US-0159584
PR	21-OCT-1999	99US-0160741
PR	21-OCT-1999	99US-0160741
PR	22-OCT-1999	99US-0160981
PR	22-OCT-1999	99US-0160981
PR	25-OCT-1999	99US-0161404
PR	25-OCT-1999	99US-0161405
PR	26-OCT-1999	99US-0161406
PR	26-OCT-1999	99US-0161359
PR	26-OCT-1999	99US-0161360
PR	26-OCT-1999	99US-0161361
PR	28-OCT-1999	99US-0161920
PR	28-OCT-1999	99US-0161922

```

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
*Query Match ~.~ 1.2%; Score 16; DB 21; Length 346;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNN 757
DB 230 NNNNNNNNNNNNNNNN 245

RESULT 3
AA182392
ID AA182392 standard; Protein; 349 AA.
AC AA182392;
XX AA182392;
XX 27-JUN-2000 (first entry)
DE C. trachomatis MOMP containing fusion protein SEQ ID NO:5.
XX Chlamydia trachomatis; fusion protein; major outer membrane protein;
XX MOMP; hydrophilic polypeptide; antibody; detection; diagnosis;
XX infection; infectious disease.
XX Chlamydia trachomatis.
XX Synthetic.
XX JP2000041678-A.
XX 15-FEB-2000.
XX 28-JUL-1998; 98JP-0213212.
XX 28-JUL-1998; 98JP-0213212.
XX (ELED ) DENKI KAGAKU KOGYO KK.
XX WPI; 2000-295780/26.
XX N-PSDB; AAA08124.
XX A soluble fused protein useful for diagnosis of Chlamydia infection,
XX comprises at least part of major outer membrane protein (MOMP) of
XX Chlamydia trachomatis -
XX
XX PS Claim 9; Page 24-25; 37pp; Japanese.
XX
XX The present invention describes fusion proteins (I) comprising at least
XX part of a major outer membrane protein (MOMP) of Chlamydia trachomatis,
XX at least one hydrophilic polypeptide having no immunoreactivity to
XX human serum and their connected part. AAA08120 to AAA08125 encode
XX specifically claimed examples of the fusion proteins given in AA182388
XX to AA182393. Also described is a method (A) for the detection of
XX Chlamydia trachomatis antibody using (I) as the antigen. (I) is used for
XX the diagnosis of Chlamydia trachomatis infectious diseases. The method
XX can diagnose Chlamydia trachomatis infectious diseases specifically in a
XX high sensitivity.
XX
XX Sequence 349 AA;
SQ

Query Match 1.2%; Score 16; DB 21; Length 349;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNN 757
DB 168 NNNNNNNNNNNNNNNN 183

RESULT 34
AA019019
ID AA019019 standard; Protein; 357 AA

```


XX AA019019;
 AC 16-DEC-2002 (first entry)
 DT
 XX Human kinesin motor protein HsKif16a motor domain.
 XX
 KW Human; kinesin motor protein; kinesin; HsKif16a; cancer; motor domain;
 KW vesicular transport disorder; neurological disorder; immune disorder;
 KW inflammation; cytoskeletal; neuroprotective; gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 16 /note= "encoded by AGN"
 FT Misc-difference 17 /note= "encoded by NNN"
 FT Misc-difference 18 /note= "encoded by NNN"
 FT Misc-difference 19 /note= "encoded by NNN"
 FT Misc-difference 20 /note= "encoded by NNN"
 FT Misc-difference 21 /note= "encoded by NNN"
 FT Misc-difference 22 /note= "encoded by NNN"
 FT Misc-difference 23 /note= "encoded by NNN"
 FT Misc-difference 24 /note= "encoded by NNN"
 FT Misc-difference 25 /note= "encoded by NNN"
 FT Misc-difference 26 /note= "encoded by NNN"
 FT Misc-difference 27 /note= "encoded by NNN"
 FT Misc-difference 28 /note= "encoded by NNN"
 FT Misc-difference 29 /note= "encoded by NNN"
 FT Misc-difference 30 /note= "encoded by NNN"
 FT Misc-difference 31 /note= "encoded by NNN"
 FT Misc-difference 32 /note= "encoded by NNN"
 FT Misc-difference 33 /note= "encoded by NNN"
 FT Misc-difference 34 /note= "encoded by NNN"
 FT Misc-difference 35 /note= "encoded by NNN"
 FT Misc-difference 36 /note= "encoded by NNN"
 FT Misc-difference 37 /note= "encoded by NNN"
 FT
 XX US6420162-B1.
 PN
 XX 16-JUL-2002.
 PD
 XX 22-NOV-2000; 2000US-0718810.
 PF
 XX 22-NOV-2000; 2000US-0718810.
 PR
 XX (CYTO-) CYTOKINETICS INC.
 PA
 XX Beraud C, Freedman R;
 PI
 XX WPI; 2002-664560/71.
 DR
 N-PSDB; AAL49367.

XX New nucleic acid encoding a motor protein which has microtubule
 PT stimulated ATPase activity, designated HsKif16a, for preventing and
 PT treating cancer, and neurological and vesicle transport disorders -
 XX
 XX Claim 1; Column 35-36; 27pp; English.
 PS
 XX The present invention provides the protein and coding sequences of a
 CC human motor protein which has microtubule stimulated ATPase activity and
 CC is designated HsKif16a. The nucleic acid is used to diagnose, treat and
 CC prevent cancer, immune disorders, inflammation, neurological disorders
 CC and disorders of vesicular transport. The present sequence is the motor
 CC domain of the protein of the invention.
 XX
 SQ Sequence 357 AA;
 QY
 Db 92 FAYGQTSGSKTYTMLG 107
 98 FAYGQTSGSKTYTMLG 113
 Query Match 1.24; Score 16; DB 23; length 357;
 Best Local Similarity 100.0%; Pred. No. 26-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 35
 AAU74558
 ID AAU74558 standard; Protein; 357 AA.
 AC
 XX AAU74558;
 DT 08-MAY-2002 (first entry)
 XX
 DE Human kinesin motor protein HsKif16a motor domain.
 XX
 KW Human; kinesin motor protein; HsKif16a; motor domain; inflammation;
 KW microtubule-stimulated ATPase activity; adenosine triphosphatase; cancer;
 KW neurological disorder; vesicular transport disorder; autoimmune disease;
 KW arthritis; graft rejection; proliferation; cytoskeletal; neurotropic;
 KW immunosuppressive; antarthritic; antiinflammatory.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 16 /label= unknown
 FT Misc-difference 17 /note= "Encoded by AGN"
 FT Misc-difference 17..37 /label= unknown
 FT Misc-difference 357 /note= "Encoded by 21 (NNN)"
 FT Misc-difference 357 /note= "Encoded by AT"
 FT
 XX US6333184-B1.
 PN
 XX 25-DEC-2001.
 PD
 XX 22-NOV-2000; 2000US-0718841.
 PF
 XX 22-NOV-2000; 2000US-0718841.
 PR
 XX (CYTO-) CYTOKINETICS INC.
 PA
 XX Beraud C, Freedman R;
 PI
 XX WPI; 2002-163180/21.
 DR
 N-PSDB; ABK14094.
 XX
 PT New human kinesin motor protein, HsKif16a, having
 PT microtubule-stimulated ATPase activity, useful for identifying specific
 PT modulators for e.g. treating cancer
 XX
 PS Claim 3; Fig 4; 27pp; English.

XX The invention relates to a human kinesin motor protein, HsKif16a,
CC comprising a motor domain and having microtubule-stimulated ATPase
CC (adenosine triphosphatase) activity. The protein of the invention and its
CC associated nucleic acid are useful in diagnosis, treatment and prevention
CC of cancer, neurological disorders, disorders of vesicular transport,
CC autoimmune diseases, arthritis, graft rejection, inflammation and
CC proliferation induced after medical procedures. The protein is also used
CC to identify its specific modulators, potentially useful as therapeutic
CC agents. This sequence represents the motor domain of the HsKif16a
CC polypeptide of the invention.

XX Sequence 357 AA;

Query Match 1.2%; Score 16; DB 23; Length 357;

Best Local Similarity 100.0%; Pred. No. 2e-06; Mismatches 0; Indels 0; Gaps 0;

Qy 92 FAYGQTGSGKTYTMLG 107
|||
Db 98 FAYGQTGSGKTYTMLG 113

RESULT 36

AA616707 standard; Protein; 376 AA.

AA616707;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 17456.

XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX EPI033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147312.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 15-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160766.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.2%; Score 16; DB 21; Length 376;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNNN 757
Db 260 NNNNNNNNNNNNNNNNN 275

RESULT 37
AAG16706
ID AAG16706 standard; Protein; 378 AA.

XX AC AAG16706;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 17455.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

XX Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141827.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142290.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143524.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148341.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

AC AA019018;
 XX
 DT 16-DEC-2002 (first entry)
 XX
 DE Human kinesin motor protein HsKif16a.
 XX
 KW Human; kinesin motor protein; kinesin; HsKif16a; cancer;
 KW vesicular transport disorder; neurological disorder; immune disorder;
 KW inflammation; cytostatic; neuroprotective; gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 16 /note= "encoded by AGN"
 FT 17 /note= "encoded by NNN"
 FT 18 /note= "encoded by NNN"
 FT 19 /note= "encoded by NNN"
 FT 20 /note= "encoded by NNN"
 FT 21 /note= "encoded by NNN"
 FT 22 /note= "encoded by NNN"
 FT 23 /note= "encoded by NNN"
 FT 24 /note= "encoded by NNN"
 FT 25 /note= "encoded by NNN"
 FT 26 /note= "encoded by NNN"
 FT 27 /note= "encoded by NNN"
 FT 28 /note= "encoded by NNN"
 FT 29 /note= "encoded by NNN"
 FT 30 /note= "encoded by NNN"
 FT 31 /note= "encoded by NNN"
 FT 32 /note= "encoded by NNN"
 FT 33 /note= "encoded by NNN"
 FT 34 /note= "encoded by NNN"
 FT 35 /note= "encoded by NNN"
 FT 36 /note= "encoded by NNN"
 FT 37 /note= "encoded by NNN"
 FT 391 /note= "encoded by NNN"
 FT 392 /note= "encoded by NNN"
 FT 393 /note= "encoded by NNN"
 FT 394 /note= "encoded by NNN"
 FT 395 /note= "encoded by NNN"
 FT 396 /note= "encoded by NNN"
 FT 397 /note= "encoded by NNN"
 FT 398 /note= "encoded by NNN"
 FT /note= "encoded by NNN"

FT Misc-difference 399 /note= "encoded by NNN"
 FT Misc-difference 400 /note= "encoded by NNN"
 FT Misc-difference 401 /note= "encoded by NNN"
 FT Misc-difference 402 /note= "encoded by NNN"
 FT Misc-difference 403 /note= "encoded by NNN"
 FT Misc-difference 404 /note= "encoded by NNN"
 FT Misc-difference 405 /note= "encoded by NNN"
 FT Misc-difference 406 /note= "encoded by NNN"
 FT Misc-difference 407 /note= "encoded by NNN"
 FT Misc-difference 408 /note= "encoded by NNN"
 FT Misc-difference 409 /note= "encoded by NNN"
 FT Misc-difference 410 /note= "encoded by NNN"
 FT Misc-difference 411 /note= "encoded by NNN"
 FT Misc-difference 412 /note= "encoded by NNN"
 FT /note= "encoded by NNN"

US6420162-B1.

16-JUL-2002.

22-NOV-2000; 2000US-0718810.

22-NOV-2000; 2000US-0718810.

(CYTO-) CYTOKINETICS INC.

Beraud C, Freedman R;

WPI, 2002-664560/71.

DR N-PSDB; AAL49366.

XX New nucleic acid encoding a motor protein which has microtubule
 PT stimulated ATPase activity, designated HSKIF16A, for preventing and
 PT treating cancer, and neurological and vesicle transport disorders -
 XX Claim 1, Column 29-34; 27pp; English.

XX The present invention provides the protein and coding sequences of a
 CC human motor protein which has microtubule stimulated ATPase activity and
 CC is designated HSKIF16A. The nucleic acid is used to diagnose, treat and
 CC prevent cancer, immune disorders, inflammation, neurological disorders
 CC and disorders of vesicular transport. The present sequence is the protein
 CC of the invention.

XX Sequence 563 AA;

Query Match 1.2%; Score 16; DB 23; Length 563;
 Best Local Similarity 100.0%; Pred. No. 3e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 FAYGOTGSGKTYTMG 107
 |||||
 Db 98 FAYGOTGSGKTYTMG 113

Search completed: October 2, 2003, 16:32:15
 Job time : 115 secs

THIS PAGE BLANK (25PT0)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 2, 2003, 16:30:31 ; Search time 23 Seconds
(without alignments)
2369.408 Million cell updates/sec

Title: US-10-006-780-2

Perfect score: 1288

Sequence: 1 MNSKIKVVRKRLSELEKK.....KLVODNIKSNMNNHKK 1288

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size: 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	1.2	286	4	US-09-384-162-6
2	16	1.2	357	4	US-09-718-841-4
3	16	1.2	357	4	US-09-718-810-4
4	16	1.2	428	3	US-09-118-319-5
5	16	1.2	432	3	US-09-118-319-2
6	16	1.2	563	4	US-09-718-841-2
7	16	1.2	563	4	US-09-718-810-2
8	16	1.2	732	4	US-08-914-999-8
9	16	1.2	1311	1	US-08-340-011-5
10	16	1.2	1311	3	US-08-901-710-5
11	16	1.2	1584	4	US-09-457-040B-27
12	16	1.2	2184	4	US-09-417-485D-6
13	15	1.2	303	4	US-09-722-129-4
14	15	1.2	492	4	US-09-722-129-2
15	15	1.2	794	4	US-09-417-485D-8
16	15	1.2	888	2	US-08-861-464-6
17	15	1.2	888	2	US-08-396-001-6
18	15	1.2	888	3	US-09-323-433A-6
19	15	1.1	23	3	US-09-098-901-11
20	14	1.1	337	4	US-09-641-806-4
21	14	1.1	337	4	US-09-723-129-4
22	14	1.1	337	4	US-09-722-862-4
23	14	1.1	342	4	US-09-641-806-2
24	14	1.1	342	4	US-09-723-129-2
25	14	1.1	342	4	US-09-722-862-2
26	14	1.1	730	4	US-09-291-170A-3
27	14	1.1	730	4	US-09-724-884-3

28	14	1.1	955	1	US-08-006-676B-1	Sequence 1, Appl1
29	14	1.1	955	1	US-08-282-845-2	Sequence 2, Appl1
30	14	1.1	955	1	US-08-428-41A-3	Sequence 3, Appl1
31	14	1.1	955	5	PCT-US94-00324-1	Sequence 1, Appl1
32	13	1.0	341	4	US-09-724-517-4	Sequence 4, Appl1
33	13	1.0	341	4	US-09-641-807A-4	Sequence 4, Appl1
34	13	1.0	341	4	US-09-723-096-4	Sequence 4, Appl1
35	13	1.0	1151	4	US-09-177-165A-31	Sequence 31, Appl1
36	13	1.0	1279	4	US-09-724-517-2	Sequence 2, Appl1
37	13	1.0	1279	4	US-09-641-807A-2	Sequence 2, Appl1
38	13	1.0	1279	4	US-09-723-096-2	Sequence 2, Appl1
39	12	0.9	246	3	US-09-451-117-2	Sequence 2, Appl1
40	12	0.9	246	4	US-09-888-655-2	Sequence 2, Appl1
41	12	0.9	362	4	US-09-594-669-6	Sequence 6, Appl1
42	12	0.9	391	4	US-09-594-669-4	Sequence 4, Appl1
43	12	0.9	405	4	US-09-594-669-12	Sequence 12, Appl1
44	12	0.9	430	4	US-09-594-669-2	Sequence 2, Appl1
45	12	0.9	434	4	US-09-594-669-10	Sequence 10, Appl1
46	12	0.9	473	4	US-09-594-669-8	Sequence 8, Appl1
47	12	0.9	475	4	US-08-386-495-10	Sequence 10, Appl1
48	12	0.9	675	5	PCT-US96-02331-10	Sequence 10, Appl1
49	12	0.9	678	4	US-09-595-684B-25	Sequence 25, Appl1
50	12	0.9	723	4	US-09-594-669-14	Sequence 14, Appl1
51	12	0.9	725	4	US-09-594-669-16	Sequence 16, Appl1
52	12	0.9	725	4	US-09-595-684B-33	Sequence 33, Appl1
53	12	0.9	842	5	PCT-US96-02331-15	Sequence 15, Appl1
54	11	0.9	205	4	US-09-637-481-2	Sequence 2, Appl1
55	11	0.9	205	4	US-09-723-428-2	Sequence 2, Appl1
56	11	0.9	205	4	US-09-724-520-2	Sequence 2, Appl1
57	11	0.9	319	4	US-09-724-510-2	Sequence 2, Appl1
58	11	0.9	319	4	US-09-723-216-2	Sequence 2, Appl1
59	11	0.9	319	4	US-09-675-227-2	Sequence 2, Appl1
60	11	0.9	343	4	US-09-967-908A-8	Sequence 8, Appl1
61	11	0.9	361	4	US-09-596-541-2	Sequence 2, Appl1
62	11	0.9	361	4	US-09-723-595-2	Sequence 2, Appl1
63	11	0.9	363	4	US-09-967-908A-10	Sequence 10, Appl1
64	11	0.9	369	4	US-09-596-541-6	Sequence 6, Appl1
65	11	0.9	369	4	US-09-723-595-6	Sequence 6, Appl1
66	11	0.9	370	4	US-09-596-541-4	Sequence 4, Appl1
67	11	0.9	370	4	US-09-723-595-4	Sequence 4, Appl1
68	11	0.9	373	4	US-09-967-908A-4	Sequence 4, Appl1
69	11	0.9	375	4	US-09-572-191-4	Sequence 4, Appl1
70	11	0.9	375	4	US-09-723-219-4	Sequence 4, Appl1
71	11	0.9	375	4	US-09-723-219-4	Sequence 4, Appl1
72	11	0.9	381	4	US-09-967-908A-6	Sequence 6, Appl1
73	11	0.9	409	4	US-09-572-191-6	Sequence 6, Appl1
74	11	0.9	409	4	US-09-723-262-6	Sequence 6, Appl1
75	11	0.9	409	4	US-09-723-219-6	Sequence 6, Appl1
76	11	0.9	473	4	US-09-592-054-6	Sequence 6, Appl1
77	11	0.9	519	4	US-09-595-684B-37	Sequence 37, Appl1
78	11	0.9	522	4	US-09-592-054-4	Sequence 4, Appl1
79	11	0.9	935	4	US-09-914-259-25	Sequence 25, Appl1
80	11	0.9	1231	4	US-09-595-684B-23	Sequence 23, Appl1
81	11	0.9	1232	4	US-09-592-054-2	Sequence 2, Appl1
82	11	0.9	1234	4	US-09-592-054-8	Sequence 8, Appl1
83	11	0.9	1368	4	US-09-967-908A-2	Sequence 2, Appl1
84	11	0.9	1388	4	US-09-572-191-2	Sequence 2, Appl1
85	11	0.9	1388	4	US-09-723-262-2	Sequence 2, Appl1
86	11	0.9	1388	4	US-09-723-219-2	Sequence 2, Appl1
87	10	0.8	24	3	US-09-098-901-9	Sequence 9, Appl1
88	10	0.8	40	1	US-08-478-312-54	Sequence 54, Appl1
89	10	0.8	40	1	US-08-485-302-54	Sequence 54, Appl1
90	10	0.8	40	1	US-08-476-169-59	Sequence 59, Appl1
91	10	0.8	40	1	US-08-484-083-59	Sequence 59, Appl1
92	10	0.8	40	2	US-08-764-640-238	Sequence 238, Appl1
93	10	0.8	40	3	US-08-973-225-226	Sequence 226, Appl1
94	10	0.8	40	3	US-09-244-296A-238	Sequence 238, Appl1
95	10	0.8	40	3	US-09-516-704-238	Sequence 238, Appl1
96	10	0.8	40	4	US-09-549-090-226	Sequence 226, Appl1
97	10	0.8	40	4	US-09-832-230A-238	Sequence 238, Appl1
98	10	0.8	100	2	US-08-710-249-16	Sequence 16, Appl1
99	10	0.8	100	4	US-09-220-157A-16	Sequence 16, Appl1
100	10	0.8	274	4	US-09-723-430-2	Sequence 2, Appl1

101	10	0.8	274	4	US-09-632-155-2	Sequence 2, Appli	174	8	0.6	444	1	US-07-881-075-3	Sequence 3, Appli
102	10	0.8	274	4	US-09-724-518-2	Sequence 2, Appli	175	8	0.6	444	1	US-08-120-827-3	Sequence 3, Appli
103	10	0.8	359	4	US-09-722-139-4	Sequence 4, Appli	176	8	0.6	444	1	US-08-478-675-3	Sequence 3, Appli
104	10	0.8	359	4	US-09-721-832-4	Sequence 4, Appli	177	8	0.6	446	1	US-09-004-3938-4	Sequence 4, Appli
105	10	0.8	359	4	US-09-721-889-4	Sequence 4, Appli	178	8	0.6	441	1	US-08-186-222-2	Sequence 2, Appli
106	10	0.8	412	4	US-09-471-396-1	Sequence 1, Appli	179	8	0.6	481	3	US-08-155-888-2	Sequence 2, Appli
107	10	0.8	412	4	US-09-788-345-12	Sequence 12, Appli	180	8	0.6	554	3	US-09-319-989-6	Sequence 6, Appli
108	10	0.8	522	2	US-08-164-614A-10	Sequence 10, Appli	181	8	0.6	586	6	US-08-419-810-12	Sequence 12, Appli
109	10	0.8	522	2	US-08-456-489B-10	Sequence 10, Appli	182	8	0.6	5405775-11	Patent No. 5405775		
110	10	0.8	536	2	US-08-164-614A-12	Sequence 12, Appli	183	8	0.6	652	4	US-08-559-896B-2	Sequence 2, Appli
111	10	0.8	536	2	US-08-456-489B-12	Sequence 12, Appli	184	8	0.6	653	4	US-09-513-057C-13	Sequence 13, Appli
112	10	0.8	596	4	US-09-752-165-2	Sequence 2, Appli	185	8	0.6	673	4	US-09-107-532A-5134	Sequence 14, Appli
113	10	0.8	659	2	US-08-709-874A-20	Sequence 20, Appli	186	8	0.6	674	4	US-08-653-648A-14	Sequence 3, Appli
114	10	0.8	659	2	US-09-104-382-20	Sequence 20, Appli	187	8	0.6	678	5	PCT-US93-03027-3	Sequence 3, Appli
115	10	0.8	675	2	US-08-971-036-2	Sequence 2, Appli	188	8	0.6	787	4	US-09-721-383-2	Sequence 2, Appli
116	10	0.8	675	2	US-09-036-570-2	Sequence 2, Appli	189	8	0.6	787	4	US-09-721-137-2	Sequence 2, Appli
117	10	0.8	675	2	US-09-265-617B-2	Sequence 2, Appli	190	8	0.6	834	2	US-08-861-464-4	Sequence 4, Appli
118	10	0.8	694	3	US-08-816-977-47	Sequence 47, Appli	191	8	0.6	834	2	US-08-396-001-1	Sequence 4, Appli
119	10	0.8	694	3	US-08-816-977-49	Sequence 49, Appli	192	8	0.6	834	3	US-09-323-333A-4	Sequence 4, Appli
120	10	0.8	708	3	US-08-816-977-33	Sequence 33, Appli	193	8	0.6	984	1	US-08-257-073-3	Sequence 3, Appli
121	10	0.8	711	3	US-08-816-977-35	Sequence 35, Appli	194	8	0.6	984	2	US-08-184-009-120	Sequence 3, Appli
122	10	0.8	809	4	US-09-650-352-5	Sequence 5, Appli	195	8	0.6	984	2	US-08-458-356-120	Sequence 120, Appli
123	10	0.8	1084	1	US-08-717-515-6	Sequence 6, Appli	196	8	0.6	984	3	US-08-460-336-120	Sequence 120, Appli
124	10	0.8	1276	1	US-08-717-515-8	Sequence 8, Appli	197	8	0.6	984	4	US-09-535-370-120	Sequence 120, Appli
125	10	0.8	1375	4	US-09-722-139-2	Sequence 2, Appli	198	8	0.6	1242	4	US-09-508-691-1	Sequence 2, Appli
126	10	0.8	1375	4	US-09-721-832-2	Sequence 2, Appli	199	8	0.6	1243	2	US-08-557-139-2	Sequence 2, Appli
127	10	0.8	1375	4	US-09-721-689-2	Sequence 2, Appli	200	8	0.6	1284	4	US-09-343-494-9	Sequence 9, Appli
128	9	0.7	11	1	US-08-040-548-43	Sequence 43, Appli	201	8	0.6	1284	4	US-09-358-383C-11	Sequence 11, Appli
129	9	0.7	11	1	US-08-466-344-43	Sequence 43, Appli	202	8	0.6	1341	2	US-08-317-310A-64	Sequence 64, Appli
130	9	0.7	11	6	5206152-12	Patent No. 5206152	203	8	0.6	1850	4	US-09-620-093A-5	Sequence 5, Appli
131	9	0.7	23	3	US-09-098-901-7	Sequence 7, Appli	204	8	0.6	2314	2	US-09-268-347-49	Sequence 49, Appli
132	9	0.7	177	4	US-09-247-155-164	Sequence 164, App	205	8	0.6	2331	2	US-08-446-855A-2	Sequence 2, Appli
133	9	0.7	330	4	US-09-134-001C-4002	Sequence 4002, Ap	206	8	0.6	2391	3	US-09-150-741-2	Sequence 2, Appli
134	9	0.7	355	4	US-09-724-511-4	Sequence 4, Appli	207	7	0.5	9	1	US-08-106-981-11	Sequence 11, Appli
135	9	0.7	355	4	US-09-723-097-4	Sequence 4, Appli	208	7	0.5	14	1	US-08-346-293-17	Sequence 17, Appli
136	9	0.7	355	4	US-09-632-344-4	Sequence 4, Appli	209	7	0.5	14	2	US-08-727-688-23	Sequence 23, Appli
137	9	0.7	367	4	US-09-724-511-2	Sequence 2, Appli	210	7	0.5	14	4	US-08-248-058-1	Sequence 1, Appli
138	9	0.7	367	4	US-09-723-097-2	Sequence 2, Appli	211	7	0.5	24	2	US-08-752-891-7	Sequence 7, Appli
139	9	0.7	367	4	US-09-632-344-2	Sequence 2, Appli	212	7	0.5	24	2	US-09-144-178-7	Sequence 7, Appli
140	9	0.7	416	4	US-09-457-040B-10	Sequence 10, Appli	213	7	0.5	24	3	US-09-406-854-7	Sequence 7, Appli
141	9	0.7	677	4	US-09-328-352-4365	Sequence 4365, Ap	214	7	0.5	43	4	US-09-098-801-6	Sequence 6, Appli
142	9	0.7	1003	1	US-08-571-758-4	Sequence 4, Appli	215	7	0.5	43	4	US-07-757-022B-12	Sequence 12, Appli
143	9	0.7	1003	1	US-08-909-984A-4	Sequence 4, Appli	216	7	0.5	51	4	US-09-314-268-100	Sequence 100, Appli
144	9	0.7	1003	1	US-08-909-983-4	Sequence 4, Appli	217	7	0.5	69	2	US-08-583-669-1	Sequence 1, Appli
145	9	0.7	2938	5	PCT-US94-00198-3	Sequence 3, Appli	218	7	0.5	73	4	US-09-328-352-4164	Sequence 4164, Ap
146	8	0.6	90	4	US-09-186-276B-59	Sequence 59, Appli	219	7	0.5	75	3	US-08-927-219-51	Sequence 51, Appli
147	8	0.6	90	4	US-08-842-445-59	Sequence 59, Appli	220	7	0.5	85	4	US-09-134-001C-3387	Sequence 3387, Ap
148	8	0.6	90	4	US-09-186-188B-59	Sequence 59, Appli	221	7	0.5	101	2	US-09-858-664A-10	Sequence 10, Appli
149	8	0.6	193	2	US-08-679-765-3	Sequence 3, Appli	222	7	0.5	110	2	US-08-569-166-34	Sequence 34, Appli
150	8	0.6	193	2	US-09-196-525-3	Sequence 3, Appli	223	7	0.5	110	4	US-08-961-083-102	Sequence 102, App
151	8	0.6	193	2	US-09-318-317-3	Sequence 3, Appli	224	7	0.5	110	4	US-09-536-784-102	Sequence 102, App
152	8	0.6	261	6	532058-5	Patent No. 532058	225	7	0.5	111	4	US-09-858-664A-8	Sequence 8, Appli
153	8	0.6	265	4	US-09-611-659A-4	Sequence 4, Appli	226	7	0.5	129	4	US-09-328-352-6408	Sequence 6408, Ap
154	8	0.6	272	2	US-08-492-027A-4	Sequence 4, Appli	227	7	0.5	130	2	US-08-630-822A-90	Sequence 90, Appli
155	8	0.6	299	3	US-09-621-233-2	Sequence 2, Appli	228	7	0.5	130	2	US-09-005-069-90	Sequence 90, Appli
156	8	0.6	299	3	US-09-724-508-2	Sequence 2, Appli	229	7	0.5	130	4	US-09-171-156A-39	Sequence 39, Appli
157	8	0.6	299	4	US-09-724-516-2	Sequence 2, Appli	230	7	0.5	130	4	US-09-004-730A-39	Sequence 39, Appli
158	8	0.6	301	4	US-09-134-001C-2915	Sequence 2915, Ap	231	7	0.5	130	4	US-08-981-799A-39	Sequence 39, Appli
159	8	0.6	317	1	US-08-221-750A-13	Sequence 13, Appli	232	7	0.5	148	4	US-09-134-001C-4679	Sequence 4679, Ap
160	8	0.6	324	4	US-09-183-861-53	Sequence 53, Appli	233	7	0.5	148	4	US-09-461-325-453	Sequence 43, App
161	8	0.6	324	4	US-09-022-765-53	Sequence 53, Appli	234	7	0.5	159	3	US-08-992-176-6	Sequence 6, Appli
162	8	0.6	324	4	US-09-551-974A-53	Sequence 53, Appli	235	7	0.5	159	4	US-09-107-532A-6434	Sequence 6434, Ap
163	8	0.6	326	4	US-09-721-383-4	Sequence 4, Appli	236	7	0.5	172	4	US-07-757-022B-88	Sequence 88, Appli
164	8	0.6	326	4	US-09-721-177-4	Sequence 4, Appli	237	7	0.5	179	4	US-09-252-891A-21165	Sequence 21165, A
165	8	0.6	342	1	US-09-118-464-3	Sequence 3, Appli	238	7	0.5	179	4	US-09-198-452A-1064	Sequence 1064, Ap
166	8	0.6	361	4	US-08-415-751-3	Sequence 3, Appli	239	7	0.5	180	4	US-09-134-001C-2966	Sequence 2966, Ap
167	8	0.6	361	4	US-09-404-296B-30	Sequence 30, Appli	240	7	0.5	187	4	US-09-247-155-102	Sequence 102, Appli
168	8	0.6	372	4	US-09-118-464-2	Sequence 2, Appli	241	7	0.5	188	3	US-09-130-663-30	Sequence 30, Appli
169	8	0.6	406	1	US-07-973-431B-1	Sequence 1, Appli	242	7	0.5	188	3	US-09-332-334-14	Sequence 14, Appli
170	8	0.6	413	3	US-08-481-814A-8	Sequence 8, Appli	243	7	0.5	188	3	US-09-432-335-30	Sequence 30, Appli
171	8	0.6	413	3	US-08-836-582-2	Sequence 2, Appli	244	7	0.5	188	3	US-09-614-022-30	Sequence 30, Appli
172	8	0.6	413	4	US-09-265-566-2	Sequence 2, Appli	245	7	0.5	192	4	US-07-757-022B-90	Sequence 90, Appli
173	8	0.6	413	4	US-09-242-737-4	Sequence 4, Appli	246	7	0.5	204	4	US-07-757-022B-92	Sequence 92, Appli

247	7	0.5	206	3	US-09-156-580-2	Sequence 2, Appl1	320	7	0.5	377	1	US-08-525-697-2	Sequence 2, Appl1
248	7	0.5	207	4	US-09-199-637A-211	Sequence 211, App	321	7	0.5	382	4	US-08-213-419B-19	Sequence 19, Appl1
249	7	0.5	208	4	US-07-757-022B-132	Sequence 132, App	322	7	0.5	382	4	US-09-724-519-8	Sequence 4, Appl1
250	7	0.5	209	4	US-07-757-022B-94	Sequence 94, Appl	323	7	0.5	382	4	US-09-592-037-4	Sequence 4, Appl1
251	7	0.5	211	2	US-08-708-958-2	Sequence 2, Appl1	324	7	0.5	391	1	US-08-602-010A-6	Sequence 6, Appl1
252	7	0.5	211	4	US-09-502-769-10	Sequence 10, Appl	325	7	0.5	391	1	US-08-680-726A-6	Sequence 6, Appl1
253	7	0.5	211	4	US-09-972-800A-12	Sequence 12, Appl	326	7	0.5	393	3	US-09-032-409-6	Sequence 6, Appl1
254	7	0.5	212	4	US-09-302-769-4	Sequence 4, Appl1	327	7	0.5	393	4	US-09-230-371A-29	Sequence 29, Appl1
255	7	0.5	212	4	US-09-392-769-12	Sequence 12, Appl	328	7	0.5	413	2	US-08-282-197C-19	Sequence 49, Appl1
256	7	0.5	212	4	US-09-399-913-63	Sequence 63, Appl	329	7	0.5	414	1	US-07-667-276A-4	Sequence 4, Appl1
257	7	0.5	212	4	US-09-341-349D-2	Sequence 2, Appl1	330	7	0.5	422	3	US-08-872-978-8	Sequence 8, Appl1
258	7	0.5	217	4	US-07-757-022B-76	Sequence 76, Appl	331	7	0.5	422	4	US-07-757-022B-68	Sequence 68, Appl
259	7	0.5	218	1	US-08-463-115-92	Sequence 92, Appl	332	7	0.5	423	4	US-07-757-022B-66	Sequence 66, Appl
260	7	0.5	218	3	US-08-465-388-92	Sequence 9, Appl1	333	7	0.5	433	4	US-09-595-424-2	Sequence 2, Appl1
261	7	0.5	218	3	US-09-068-655-7	Sequence 7, Appl1	334	7	0.5	446	4	US-09-286-981B-6	Sequence 6, Appl1
262	7	0.5	219	1	US-08-463-115-91	Sequence 91, Appl	335	7	0.5	446	4	US-09-286-981B-9	Sequence 9, Appl1
263	7	0.5	219	1	US-08-465-388-91	Sequence 91, Appl	336	7	0.5	450	2	US-08-861-464-2	Sequence 2, Appl1
264	7	0.5	219	4	US-09-134-001C-4581	Sequence 4581, Ap	337	7	0.5	450	2	US-08-396-001-2	Sequence 2, Appl1
265	7	0.5	220	4	US-07-757-022B-96	Sequence 96, Appl1	338	7	0.5	450	3	US-09-323-433A-2	Sequence 2, Appl1
266	7	0.5	223	4	US-09-206-676C-1	Sequence 1, Appl1	339	7	0.5	460	4	US-09-134-001C-5322	Sequence 5322, Ap
267	7	0.5	223	4	US-09-009-816-4	Sequence 4, Appl1	340	7	0.5	463	4	US-07-757-022B-54	Sequence 54, Appl
268	7	0.5	226	4	US-09-206-676C-2	Sequence 2, Appl1	341	7	0.5	463	4	US-09-345-236B-2	Sequence 2, Appl1
269	7	0.5	229	4	US-09-604-978-9	Sequence 9, Appl1	342	7	0.5	465	3	US-09-537-357-52	Sequence 52, Appl
270	7	0.5	229	4	US-09-604-978-9	Sequence 9, Appl1	343	7	0.5	465	4	US-09-252-991A-17635	Sequence 17635, A
271	7	0.5	230	3	US-09-248-335-44	Sequence 44, Appl	344	7	0.5	467	4	US-08-495-484-12	Sequence 12, Appl1
272	7	0.5	231	4	US-07-757-022B-30	Sequence 30, Appl	345	7	0.5	471	1	US-08-176-427B-11	Sequence 11, Appl
273	7	0.5	233	2	US-08-458-568A-4	Sequence 4, Appl1	346	7	0.5	471	2	US-08-356-060A-34	Sequence 34, Appl
274	7	0.5	243	4	US-09-166-350-19	Sequence 19, Appl	347	7	0.5	471	3	US-08-460-900C-34	Sequence 34, Appl
275	7	0.5	246	3	US-09-185-160-7	Sequence 7, Appl1	348	7	0.5	471	3	US-08-757-230A-8	Sequence 8, Appl1
276	7	0.5	246	2	US-08-861-269-5	Sequence 5, Appl1	349	7	0.5	471	3	US-08-674-509B-34	Sequence 34, Appl1
277	7	0.5	250	2	US-09-134-596-5	Sequence 5, Appl1	350	7	0.5	471	3	US-08-954-698-34	Sequence 34, Appl
278	7	0.5	250	3	US-09-293-273-5	Sequence 5, Appl1	351	7	0.5	471	4	US-08-957-874-34	Sequence 34, Appl
279	7	0.5	254	4	US-09-294-531B-3	Sequence 3, Appl1	352	7	0.5	471	4	US-08-700-393-8	Sequence 8, Appl
280	7	0.5	258	2	US-08-720-258-4	Sequence 4, Appl1	353	7	0.5	474	4	US-09-639-699-34	Sequence 34, Appl
281	7	0.5	262	1	US-08-106-981-6	Sequence 6, Appl1	354	7	0.5	474	4	US-09-702-705-1812	Sequence 1812, Ap
282	7	0.5	264	1	US-08-463-115-93	Sequence 93, Appl	355	7	0.5	474	4	US-09-736-457-1812	Sequence 1812, Ap
283	7	0.5	264	3	US-08-465-388-93	Sequence 93, Appl	356	7	0.5	475	2	US-08-861-464-14	Sequence 14, Appl
284	7	0.5	264	4	US-08-894-731-4	Sequence 4, Appl1	357	7	0.5	475	3	US-08-396-001-14	Sequence 14, Appl
285	7	0.5	266	3	US-09-252-991A-17646	Sequence 17646, A	358	7	0.5	475	3	US-09-323-433A-14	Sequence 14, Appl
286	7	0.5	285	3	US-08-482-085B-20	Sequence 20, Appl	359	7	0.5	475	4	US-09-569-804-4	Sequence 4, Appl1
287	7	0.5	288	4	US-09-134-001C-3292	Sequence 3292, Ap	360	7	0.5	476	3	US-09-316-083-3	Sequence 3, Appl1
288	7	0.5	296	4	US-07-757-022B-70	Sequence 70, Appl	361	7	0.5	476	4	US-09-933-700-3	Sequence 3, Appl1
289	7	0.5	302	4	US-09-725-311-2	Sequence 2, Appl1	362	7	0.5	484	4	US-09-242-913B-17	Sequence 17, Appl
290	7	0.5	310	3	US-08-651-136C-22	Sequence 22, Appl	363	7	0.5	484	4	US-09-724-224-8	Sequence 8, Appl1
291	7	0.5	314	4	US-09-229-911A-22	Sequence 22, Appl	364	7	0.5	489	1	US-07-903-103-4	Sequence 4, Appl1
292	7	0.5	314	4	US-09-544-618-20	Sequence 20, Appl	365	7	0.5	489	1	US-08-044-619A-4	Sequence 4, Appl1
293	7	0.5	315	4	US-09-404-296B-8	Sequence 8, Appl1	366	7	0.5	489	1	US-08-283-911-4	Sequence 4, Appl1
294	7	0.5	320	4	US-09-134-001C-4439	Sequence 4439, Ap	367	7	0.5	489	1	US-08-245-500A-5	Sequence 5, Appl1
295	7	0.5	325	4	US-09-252-991A-27128	Sequence 27128, A	368	7	0.5	489	1	US-08-390-546-5	Sequence 5, Appl1
296	7	0.5	330	4	US-09-725-311-4	Sequence 4, Appl1	369	7	0.5	489	1	US-08-390-479A-5	Sequence 5, Appl1
297	7	0.5	338	4	US-09-634-957-4	Sequence 4, Appl1	370	7	0.5	489	1	US-08-557-393-5	Sequence 5, Appl1
298	7	0.5	338	4	US-09-723-153-4	Sequence 4, Appl1	371	7	0.5	489	1	US-08-390-516C-5	Sequence 5, Appl1
299	7	0.5	345	2	US-09-723-429-4	Sequence 4, Appl1	372	7	0.5	489	1	US-08-390-517A-5	Sequence 5, Appl1
300	7	0.5	345	4	US-08-282-197C-50	Sequence 50, Appl	373	7	0.5	489	2	US-08-801-718-5	Sequence 5, Appl1
301	7	0.5	346	4	US-09-724-224-6	Sequence 6, Appl1	374	7	0.5	489	4	US-09-170-159A-5	Sequence 5, Appl1
302	7	0.5	347	3	US-09-094-557-3	Sequence 3, Appl1	375	7	0.5	489	4	US-09-480-718-46	Sequence 46, Appl
303	7	0.5	349	4	US-09-162-524-3	Sequence 3, Appl1	376	7	0.5	490	4	US-09-336-643A-6	Sequence 6, Appl1
304	7	0.5	351	4	US-09-399-913-61	Sequence 61, Appl	377	7	0.5	494	4	US-09-595-424-4	Sequence 4, Appl1
305	7	0.5	351	4	US-09-252-991A-17990	Sequence 17990, A	378	7	0.5	496	4	US-09-350-268-2	Sequence 2, Appl1
306	7	0.5	360	3	US-08-899-437-7	Sequence 7, Appl1	379	7	0.5	496	4	US-09-558-679-2	Sequence 2, Appl1
307	7	0.5	360	3	US-09-126-121-7	Sequence 7, Appl1	380	7	0.5	506	4	US-09-661-468-19	Sequence 19, Appl
308	7	0.5	366	4	US-09-718-692-4	Sequence 4, Appl1	381	7	0.5	506	2	US-08-820-170A-19	Sequence 19, Appl
309	7	0.5	366	4	US-09-718-852-4	Sequence 4, Appl1	382	7	0.5	506	3	US-09-055-699-19	Sequence 19, Appl
310	7	0.5	366	4	US-09-718-815-4	Sequence 4, Appl1	383	7	0.5	506	3	US-09-273-565-19	Sequence 19, Appl
311	7	0.5	367	1	US-07-737-736B-4	Sequence 4, Appl1	384	7	0.5	506	4	US-09-565-536-19	Sequence 19, Appl
312	7	0.5	367	4	US-09-009-816-2	Sequence 15, Appl	385	7	0.5	506	4	US-09-661-468-19	Sequence 19, Appl
313	7	0.5	368	3	US-09-413-574-4	Sequence 4, Appl1	386	7	0.5	512	4	US-09-976-165-19	Sequence 19, Appl
314	7	0.5	368	4	US-09-724-519-10	Sequence 10, Appl	387	7	0.5	512	4	US-09-724-519-6	Sequence 6, Appl1
315	7	0.5	368	4	US-09-592-037-10	Sequence 10, Appl	388	7	0.5	513	4	US-09-592-037-6	Sequence 6, Appl1
316	7	0.5	369	2	US-08-773-870-4	Sequence 4, Appl1	389	7	0.5	516	1	US-08-097-829-4	Sequence 4, Appl1
317	7	0.5	370	4	US-09-724-224-2	Sequence 2, Appl1	390	7	0.5	516	1	US-08-577-403-4	Sequence 4, Appl1
318	7	0.5	372	4	US-07-757-022B-64	Sequence 64, Appl	391	7	0.5	516	1		
319	7	0.5	372	4			392	7	0.5	516	1		

393	7	0.5	524	4	US-09-186-276B-56	Sequence 56, Appl	466	7	0.5	823	4	US-09-650-855-19	Sequence 19, Appl
394	7	0.5	524	4	US-08-842-445-56	Sequence 56, Appl	467	7	0.5	845	4	US-09-198-452A-48	Sequence 48, Appl
395	7	0.5	524	4	US-09-186-188B-56	Sequence 56, Appl	468	7	0.5	846	2	US-07-728-215-33	Sequence 33, Appl
396	7	0.5	527	4	US-09-134-001C-3358	Sequence 3358, Ap	469	7	0.5	846	4	US-08-938-085A-33	Sequence 33, Appl
397	7	0.5	535	4	US-09-134-001C-3338	Sequence 3338, Ap	470	7	0.5	846	4	US-10-072-444-33	Sequence 33, Appl
398	7	0.5	536	4	US-09-463-712C-10	Sequence 10, Appl	471	7	0.5	866	2	US-08-483-101-4	Sequence 4, Appl
399	7	0.5	543	4	US-08-224-482-4	Sequence 4, Appl	472	7	0.5	878	4	US-09-735-534A-2	Sequence 2, Appl
400	7	0.5	543	2	US-08-469-412A-7	Sequence 7, Appl	473	7	0.5	898	4	US-10-060-332-2	Sequence 2, Appl
401	7	0.5	543	3	US-09-021-715-7	Sequence 7, Appl	474	7	0.5	928	4	US-09-585-858-37	Sequence 37, Appl
402	7	0.5	543	4	US-09-702-705-337	Sequence 337, App	475	7	0.5	947	4	US-09-914-259-23	Sequence 23, Appl
403	7	0.5	543	4	US-09-726-457-337	Sequence 337, App	476	7	0.5	947	4	US-09-418-780A-1	Sequence 1, Appl
404	7	0.5	548	2	US-08-469-412A-2	Sequence 2, Appl	477	7	0.5	957	4	US-09-595-684B-27	Sequence 27, Appl
405	7	0.5	548	3	US-09-021-715-2	Sequence 2, Appl	478	7	0.5	957	4	US-09-107-532A-4021	Sequence 4021, Ap
406	7	0.5	551	4	US-09-252-991A-27787	Sequence 27787, A	479	7	0.5	960	4	US-09-595-424-8	Sequence 8, Appl
407	7	0.5	553	4	US-09-242-913B-13	Sequence 13, Appl	480	7	0.5	989	4	US-08-213-419B-2	Sequence 2, Appl
408	7	0.5	568	1	US-08-330-559-30	Sequence 30, Appl	481	7	0.5	989	4	US-08-521-511C-11	Sequence 11, Appl
409	7	0.5	568	3	US-08-545-860D-30	Sequence 30, Appl	482	7	0.5	1003	4	US-09-521-511C-11	Sequence 11, Appl
410	7	0.5	568	5	PCT-US94-04496-30	Sequence 30, Appl	483	7	0.5	1049	4	US-07-757-022B-58	Sequence 58, Appl
411	7	0.5	575	4	US-09-724-519-8	Sequence 8, Appl	484	7	0.5	1050	4	US-09-428-711A-16	Sequence 16, Appl
412	7	0.5	575	4	US-09-592-037-8	Sequence 8, Appl	485	7	0.5	1051	4	US-09-428-711A-14	Sequence 14, Appl
413	7	0.5	579	4	US-09-529-279-4	Sequence 4, Appl	486	7	0.5	1053	4	US-09-724-519-2	Sequence 2, Appl
414	7	0.5	579	4	US-10-158-895-4	Sequence 4, Appl	487	7	0.5	1053	4	US-09-592-037-2	Sequence 2, Appl
415	7	0.5	582	3	US-08-906-865-3	Sequence 3, Appl	488	7	0.5	1056	4	US-09-595-684B-29	Sequence 29, Appl
416	7	0.5	582	4	US-09-428-711A-2	Sequence 2, Appl	489	7	0.5	1057	3	US-09-541-782-10	Sequence 10, Appl
417	7	0.5	582	4	US-09-129-668-3	Sequence 3, Appl	490	7	0.5	1057	4	US-09-723-820-10	Sequence 10, Appl
418	7	0.5	590	2	US-08-785-310A-5	Sequence 5, Appl	491	7	0.5	1066	3	US-09-541-782-8	Sequence 8, Appl
419	7	0.5	590	4	US-09-529-279-15	Sequence 15, Appl	492	7	0.5	1066	3	US-09-723-820-8	Sequence 8, Appl
420	7	0.5	590	4	US-10-158-895-15	Sequence 15, Appl	493	7	0.5	1073	3	US-09-541-782-6	Sequence 6, Appl
421	7	0.5	592	1	US-08-217-327-8	Sequence 8, Appl	494	7	0.5	1073	4	US-09-723-820-6	Sequence 6, Appl
422	7	0.5	594	2	US-08-785-310A-6	Sequence 6, Appl	495	7	0.5	1093	3	US-08-545-860D-55	Sequence 55, Appl
423	7	0.5	607	4	US-09-252-991A-32643	Sequence 32643, A	496	7	0.5	1093	5	PCT-US94-04496-55	Sequence 55, Appl
424	7	0.5	609	4	US-09-198-452A-579	Sequence 579, App	497	7	0.5	1103	3	US-09-162-573-1	Sequence 1, Appl
425	7	0.5	637	4	US-09-284-768A-20	Sequence 20, Appl	498	7	0.5	1103	3	US-09-467-946-1	Sequence 1, Appl
426	7	0.5	646	4	US-09-328-352-6017	Sequence 6017, Ap	499	7	0.5	1111	4	US-09-914-259-28	Sequence 28, Appl
427	7	0.5	658	4	US-09-595-424-6	Sequence 6, Appl	500	7	0.5	1140	4	US-07-757-022B-104	Sequence 104, App
428	7	0.5	660	3	US-09-058-469-18	Sequence 18, Appl	501	7	0.5	1141	1	US-08-121-365B-54	Sequence 54, Appl
429	7	0.5	660	3	US-09-058-469-91	Sequence 91, Appl	502	7	0.5	1141	2	US-08-668-123-54	Sequence 54, Appl
430	7	0.5	662	3	US-09-058-489-15	Sequence 15, Appl	503	7	0.5	1142	2	US-08-993-118-7	Sequence 7, Appl
431	7	0.5	662	3	US-09-058-489-16	Sequence 16, Appl	504	7	0.5	1142	3	US-08-845-528C-7	Sequence 7, Appl
432	7	0.5	665	4	US-09-595-684B-35	Sequence 35, Appl	505	7	0.5	1142	3	US-09-061-709-2	Sequence 2, Appl
433	7	0.5	674	4	US-09-284-768A-21	Sequence 21, Appl	506	7	0.5	1142	4	US-09-066-281B-7	Sequence 7, Appl
434	7	0.5	693	4	US-09-107-532A-5812	Sequence 5812, Ap	507	7	0.5	1187	1	US-09-899-651-2	Sequence 2, Appl
435	7	0.5	694	3	US-08-559-397A-31	Sequence 31, Appl	508	7	0.5	1184	3	US-09-541-782-2	Sequence 2, Appl
436	7	0.5	696	3	US-08-899-437-23	Sequence 23, Appl	509	7	0.5	1184	4	US-09-723-820-2	Sequence 2, Appl
437	7	0.5	703	3	US-09-126-121-23	Sequence 23, Appl	510	7	0.5	1185	3	US-09-266-225D-18	Sequence 18, Appl
438	7	0.5	706	3	US-08-910-925-4	Sequence 2, Appl	511	7	0.5	1185	3	US-09-041-886-23	Sequence 23, Appl
439	7	0.5	706	3	US-09-408-820-2	Sequence 2, Appl	512	7	0.5	1187	1	US-08-320-859-28	Sequence 28, Appl
440	7	0.5	706	4	US-09-914-259-29	Sequence 29, Appl	513	7	0.5	1187	3	US-08-545-860D-28	Sequence 28, Appl
441	7	0.5	715	3	US-08-669-286-5	Sequence 5, Appl	514	7	0.5	1187	5	PCT-US94-04496-28	Sequence 28, Appl
442	7	0.5	715	3	US-09-469-253-5	Sequence 5, Appl	515	7	0.5	1210	1	US-08-320-859-26	Sequence 26, Appl
443	7	0.5	715	3	US-09-642-146-5	Sequence 5, Appl	516	7	0.5	1210	3	US-08-545-860D-26	Sequence 26, Appl
444	7	0.5	716	4	US-09-107-532A-5208	Sequence 5208, Ap	517	7	0.5	1210	5	PCT-US94-04496-26	Sequence 26, Appl
445	7	0.5	717	3	US-08-910-925-1	Sequence 1, Appl	518	7	0.5	1225	1	US-08-118-101A-2	Sequence 2, Appl
446	7	0.5	720	3	US-08-899-437-6	Sequence 6, Appl	519	7	0.5	1298	2	US-08-650-473-2	Sequence 2, Appl
447	7	0.5	720	3	US-09-126-121-6	Sequence 6, Appl	520	7	0.5	1298	3	US-09-259-821A-2	Sequence 2, Appl
448	7	0.5	731	3	US-09-185-160-11	Sequence 11, Appl	521	7	0.5	1298	3	US-08-843-659-2	Sequence 2, Appl
449	7	0.5	734	3	US-09-185-160-13	Sequence 13, Appl	522	7	0.5	1313	4	US-07-757-022B-142	Sequence 142, App
450	7	0.5	739	4	US-09-1134-001C-3586	Sequence 3586, Ap	523	7	0.5	1314	4	US-07-757-022B-58	Sequence 58, Appl
451	7	0.5	743	3	US-08-910-925-3	Sequence 3, Appl	524	7	0.5	1328	3	US-08-781-891-76	Sequence 76, Appl
452	7	0.5	750	3	US-09-185-160-14	Sequence 14, Appl	525	7	0.5	1328	4	US-09-618-166-76	Sequence 76, Appl
453	7	0.5	752	1	US-08-244-189-2	Sequence 2, Appl	526	7	0.5	1354	4	US-07-757-022B-48	Sequence 48, Appl
454	7	0.5	764	4	US-09-370-838-67	Sequence 67, Appl	527	7	0.5	1363	4	US-07-757-022B-52	Sequence 52, Appl
455	7	0.5	769	3	US-08-727-308-1	Sequence 1, Appl	528	7	0.5	1404	4	US-07-757-022B-2	Sequence 2, Appl
456	7	0.5	789	4	US-09-390-234-16	Sequence 16, Appl	529	7	0.5	1404	4	US-07-757-022B-62	Sequence 62, Appl
457	7	0.5	790	4	US-09-634-957-2	Sequence 2, Appl	530	7	0.5	1451	4	US-09-060-299-25	Sequence 25, Appl
458	7	0.5	790	4	US-09-723-153-2	Sequence 2, Appl	531	7	0.5	1451	4	US-09-402-823A-25	Sequence 25, Appl
459	7	0.5	790	4	US-09-723-429-2	Sequence 2, Appl	532	7	0.5	1481	2	US-08-616-644-40	Sequence 40, Appl
460	7	0.5	805	4	US-09-425-335-6	Sequence 6, Appl	533	7	0.5	1481	2	US-08-599-654-40	Sequence 40, Appl
461	7	0.5	809	4	US-09-186-276B-58	Sequence 58, Appl	534	7	0.5	1481	3	US-08-944-688A-40	Sequence 40, Appl
462	7	0.5	809	4	US-08-843-445-58	Sequence 58, Appl	535	7	0.5	1481	3	US-08-944-423A-40	Sequence 40, Appl
463	7	0.5	809	4	US-09-186-188B-58	Sequence 58, Appl	536	7	0.5	1481	3	US-08-944-496-40	Sequence 40, Appl
464	7	0.5	821	1	US-09-377-465A-2	Sequence 2, Appl	537	7	0.5	1581	3	US-09-110-517-2	Sequence 2, Appl
465	7	0.5	823	4	US-09-651-656-19	Sequence 19, Appl	538	7	0.5	1584	4	US-09-060-299-39	Sequence 39, Appl

539	7	0.5	1584	4	US-09-402-923A-39	Sequence 39, Appl	612	6	0.5	22	3	US-08-940-095-173	Sequence 173, App
540	7	0.5	1586	4	US-09-060-299-44	Sequence 44, Appl	613	6	0.5	22	3	US-08-940-095-175	Sequence 175, App
541	7	0.5	1586	4	US-09-402-923A-44	Sequence 44, Appl	614	6	0.5	22	3	US-08-940-093-173	Sequence 173, App
542	7	0.5	1591	4	US-09-060-299-4	Sequence 4, Appl1	615	6	0.5 ⁴	22	3	US-08-940-093-175	Sequence 175, App
544	7	0.5	1591	4	US-09-060-299-43	Sequence 43, Appl	616	6	0.5	22	3	US-08-940-096-173	Sequence 173, App
545	7	0.5	1591	4	US-09-402-923A-4	Sequence 43, Appl	617	6	0.5	22	3	US-08-940-096-175	Sequence 173, App
546	7	0.5	1591	4	US-09-402-923A-43	Sequence 43, Appl	618	6	0.5	22	3	US-09-465-719-173	Sequence 173, App
547	7	0.5	1614	4	US-09-060-299-42	Sequence 42, Appl	619	6	0.5	22	3	US-09-465-719-175	Sequence 175, App
548	7	0.5	1615	4	US-09-402-923A-42	Sequence 42, Appl	620	6	0.5	22	3	US-09-453-605-173	Sequence 173, App
549	7	0.5	1615	4	US-09-060-299-3	Sequence 3, Appl1	621	6	0.5	22	4	US-09-453-605-175	Sequence 175, App
550	7	0.5	1615	4	US-09-402-923A-3	Sequence 3, Appl1	622	6	0.5	22	4	US-09-453-838-173	Sequence 173, App
551	7	0.5	1622	4	US-09-331-899-72	Sequence 72, Appl	623	6	0.5	22	4	US-09-453-838-175	Sequence 175, App
552	7	0.5	1637	1	US-07-665-792B-9	Sequence 9, Appl1	624	6	0.5	22	4	US-08-940-136-173	Sequence 173, App
553	7	0.5	1637	4	US-09-718-692-2	Sequence 2, Appl1	625	6	0.5	22	4	US-08-940-136-175	Sequence 175, App
554	7	0.5	1637	4	US-09-718-852-2	Sequence 2, Appl1	626	6	0.5	22	4	US-09-453-841-173	Sequence 173, App
555	7	0.5	1637	4	US-09-718-815-2	Sequence 2, Appl1	627	6	0.5	22	4	US-09-453-841-175	Sequence 175, App
556	7	0.5	1639	4	US-09-060-299-29	Sequence 29, Appl	628	6	0.5	24	3	US-09-098-901-4	Sequence 4, Appl1
557	7	0.5	1639	4	US-09-402-923A-29	Sequence 29, Appl	629	6	0.5	24	3	US-09-098-901-5	Sequence 5, Appl1
558	7	0.5	1665	4	US-09-858-664A-2	Sequence 2, Appl1	630	6	0.5	24	6	US-08-473-475A-6	Patent No. 5198345
559	7	0.5	1690	4	US-09-595-684B-39	Sequence 39, Appl	631	6	0.5	25	2	US-08-473-475A-6	Sequence 6, Appl1
560	7	0.5	1706	2	US-08-459-568-2	Sequence 2, Appl1	632	6	0.5	25	3	US-09-200-757-5	Patent No. 5198345
561	7	0.5	1706	2	US-08-399-411-2	Sequence 2, Appl1	633	6	0.5	25	6	US-08-254-493-7	Sequence 7, Appl1
562	7	0.5	1706	3	US-08-516-859A-2	Sequence 2, Appl1	634	6	0.5	26	1	US-08-408-222B-7	Sequence 7, Appl1
563	7	0.5	1706	4	US-09-586-472-2	Sequence 2, Appl1	635	6	0.5	26	1	US-09-056-226-8	Sequence 8, Appl1
564	7	0.5	1719	4	US-09-528-706-2	Sequence 2, Appl1	636	6	0.5	27	3	US-09-056-226-8	Patent No. 5198345
565	7	0.5	1719	2	US-08-459-568-4	Sequence 4, Appl1	637	6	0.5	28	6	5198345-19	Sequence 39, Appl
566	7	0.5	1719	3	US-08-399-411-4	Sequence 4, Appl1	638	6	0.5	30	1	US-08-446-692-39	Sequence 39, Appl
567	7	0.5	1719	4	US-09-586-472-4	Sequence 4, Appl1	639	6	0.5	30	2	US-08-488-351A-39	Sequence 39, Appl
568	7	0.5	1719	4	US-09-528-706-4	Sequence 4, Appl1	640	6	0.5	30	3	US-09-100-414B-6	Sequence 66, Appl
569	7	0.5	1783	4	US-09-362-336A-2	Sequence 4, Appl1	641	6	0.5	30	3	US-09-100-414B-70	Sequence 70, Appl
570	7	0.5	1804	4	US-09-362-336A-4	Sequence 4, Appl1	642	6	0.5	30	3	US-09-303-323-66	Sequence 66, Appl
571	7	0.5	1805	4	US-09-004-838-92	Sequence 92, Appl	643	6	0.5	30	3	US-09-303-323-70	Sequence 70, Appl
572	7	0.5	1817	4	US-09-004-838-125	Sequence 125, App	644	6	0.5	30	4	US-09-770-014-66	Sequence 66, Appl
573	7	0.5	1854	4	US-09-004-838-108	Sequence 108, App	645	6	0.5	31	3	US-09-023-082A-25	Sequence 25, Appl
574	7	0.5	1874	4	US-09-331-403-2	Sequence 2, Appl1	646	6	0.5	34	4	US-09-301-978C-19	Sequence 19, Appl
575	7	0.5	1890	4	US-09-004-838-88	Sequence 88, Appl	647	6	0.5	34	4	US-09-082-279B-289	Sequence 289, App
576	7	0.5	2265	2	US-08-149-097D-36	Sequence 36, Appl	648	6	0.5	35	3	US-09-082-279B-290	Sequence 290, App
577	7	0.5	2509	2	US-08-149-097D-35	Sequence 35, Appl	649	6	0.5	35	3	US-09-082-279B-291	Sequence 291, App
578	7	0.5	2662	4	US-09-595-684B-31	Sequence 31, Appl	650	6	0.5	35	3	US-09-082-279B-292	Sequence 292, App
579	7	0.5	3788	4	US-09-336-447A-76	Sequence 76, Appl	651	6	0.5	35	3	US-09-082-279B-293	Sequence 293, App
580	7	0.5	3969	3	US-08-061-376-5	Sequence 5, Appl1	652	6	0.5	35	3	US-09-082-279B-294	Sequence 294, App
581	6	0.5	7	3	US-08-478-316-51	Sequence 51, Appl	653	6	0.5	35	3	US-09-082-279B-295	Sequence 295, App
582	6	0.5	7	4	US-09-019-793A-13	Sequence 13, Appl	654	6	0.5	35	3	US-09-082-279B-296	Sequence 296, App
583	6	0.5	9	4	US-09-144-280-51	Sequence 13, Appl	655	6	0.5	35	3	US-09-082-279B-297	Sequence 297, App
584	6	0.5	10	4	US-09-468-265-6	Sequence 6, Appl1	656	6	0.5	35	3	US-09-082-279B-298	Sequence 298, App
585	6	0.5	10	4	US-09-434-476A-6	Sequence 6, Appl1	657	6	0.5	35	3	US-09-082-279B-299	Sequence 299, App
586	6	0.5	10	4	US-09-434-476A-7	Sequence 7, Appl1	658	6	0.5	35	3	US-09-082-279B-300	Sequence 300, App
587	6	0.5	12	4	US-08-481-968A-27	Sequence 27, Appl	659	6	0.5	35	3	US-09-082-279B-301	Sequence 301, App
588	6	0.5	12	4	US-08-154-712B-27	Sequence 27, Appl	660	6	0.5	35	3	US-09-082-279B-302	Sequence 302, App
589	6	0.5	12	4	US-09-257-825B-26	Sequence 26, Appl	661	6	0.5	35	3	US-09-082-279B-303	Sequence 303, App
590	6	0.5	13	1	US-08-346-293-16	Sequence 16, Appl	662	6	0.5	35	3	US-09-082-279B-304	Sequence 304, App
591	6	0.5	13	1	US-08-057-167-16	Sequence 16, Appl	663	6	0.5	35	3	US-09-082-279B-305	Sequence 305, App
592	6	0.5	13	5	PCT-US93-05412-16	Sequence 16, Appl	664	6	0.5	35	3	US-09-082-279B-306	Sequence 306, App
593	6	0.5	14	6	5204097-4	Patent No. 5204097	665	6	0.5	35	4	US-09-315-304B-289	Sequence 289, App
594	6	0.5	15	4	US-08-481-968A-17	Sequence 17, Appl	666	6	0.5	35	4	US-09-315-304B-290	Sequence 290, App
595	6	0.5	15	4	US-08-154-712B-17	Sequence 17, Appl	667	6	0.5	35	4	US-09-315-304B-291	Sequence 291, App
596	6	0.5	16	3	US-08-851-843A-201	Sequence 201, App	668	6	0.5	35	4	US-09-315-304B-292	Sequence 292, App
597	6	0.5	16	3	US-08-974-549A-320	Sequence 320, App	669	6	0.5	35	4	US-09-315-304B-293	Sequence 293, App
598	6	0.5	16	3	US-08-931-858B-92	Sequence 92, Appl	670	6	0.5	35	4	US-09-315-304B-300	Sequence 300, App
599	6	0.5	16	3	US-08-981-739-92	Sequence 92, Appl	671	6	0.5	35	4	US-09-315-304B-301	Sequence 301, App
600	6	0.5	16	3	US-09-146-755-1	Sequence 1, Appl1	672	6	0.5	35	4	US-09-315-304B-302	Sequence 302, App
601	6	0.5	16	3	US-08-854-050-201	Sequence 201, App	673	6	0.5	35	4	US-09-315-304B-303	Sequence 303, App
602	6	0.5	16	4	US-09-430-323-201	Sequence 201, App	674	6	0.5	35	4	US-09-315-304B-304	Sequence 304, App
603	6	0.5	16	4	US-09-128-026-92	Sequence 26, Appl	675	6	0.5	35	4	US-09-315-304B-305	Sequence 305, App
604	6	0.5	18	3	US-09-100-414B-26	Sequence 26, Appl	676	6	0.5	35	4	US-09-315-304B-306	Sequence 306, App
605	6	0.5	18	3	US-09-100-414B-30	Sequence 30, Appl	677	6	0.5	35	4	US-09-315-304B-307	Sequence 307, App
606	6	0.5	18	3	US-09-303-323-26	Sequence 26, Appl	678	6	0.5	35	4	US-09-315-304B-308	Sequence 308, App
607	6	0.5	18	3	US-09-303-323-30	Sequence 30, Appl	679	6	0.5	35	4	US-09-315-304B-309	Sequence 309, App
608	6	0.5	18	4	US-09-770-014-26	Sequence 26, Appl	680	6	0.5	35	4	US-09-315-304B-310	Sequence 310, App
609	6	0.5	18	4	US-09-770-014-30	Sequence 30, Appl	681	6	0.5	35	4	US-09-315-304B-311	Sequence 311, App
610	6	0.5	19	3	US-08-817-926-12	Sequence 12, Appl	682	6	0.5	35	4	US-09-315-304B-312	Sequence 312, App
611	6	0.5	20	6	5204097-3	Patent No. 5204097	683	6	0.5	35	4	US-09-834-784-289	Sequence 289, App

685	6	0.5	35	4	US-09-834-784-290	Sequence 290, App	758	6	0.5	93	3	US-08-931-858E-176	Sequence 176, App
686	6	0.5	35	4	US-09-834-784-291	Sequence 291, App	759	6	0.5	93	3	US-09-220-528-16	Sequence 16, App
687	6	0.5	35	4	US-09-834-784-292	Sequence 292, App	760	6	0.5	93	3	US-08-469-260A-19	Sequence 19, App
688	6	0.5	35	4	US-09-834-784-293	Sequence 293, App	761	6	0.5	93	4	US-08-488-446-19	Sequence 19, App
689	6	0.5	35	4	US-09-834-784-294	Sequence 294, App	762	6	0.5	93	4	US-08-467-344A-19	Sequence 19, App
690	6	0.5	35	4	US-09-834-784-295	Sequence 295, App	763	6	0.5	94	3	US-08-931-858E-87	Sequence 87, App
691	6	0.5	35	4	US-09-834-784-296	Sequence 296, App	764	6	0.5	94	3	US-08-931-858E-172	Sequence 172, App
692	6	0.5	35	4	US-09-834-784-297	Sequence 297, App	765	6	0.5	94	3	US-08-981-739-87	Sequence 87, App
693	6	0.5	35	4	US-09-834-784-298	Sequence 298, App	766	6	0.5	94	3	US-08-981-739-172	Sequence 172, App
694	6	0.5	35	4	US-09-834-784-299	Sequence 299, App	767	6	0.5	94	3	US-09-128-026-87	Sequence 87, App
695	6	0.5	35	4	US-09-834-784-300	Sequence 300, App	768	6	0.5	94	4	US-09-128-026-172	Sequence 172, App
696	6	0.5	35	4	US-09-834-784-301	Sequence 301, App	769	6	0.5	97	3	US-08-816-977-23	Sequence 23, App
697	6	0.5	35	4	US-09-834-784-302	Sequence 302, App	770	6	0.5	98	4	US-09-134-001C-4659	Sequence 4659, App
698	6	0.5	35	4	US-09-834-784-303	Sequence 303, App	771	6	0.5	103	3	US-08-535-681-46	Sequence 46, App
699	6	0.5	35	4	US-09-834-784-304	Sequence 304, App	772	6	0.5	103	3	US-08-535-681-49	Sequence 49, App
700	6	0.5	35	4	US-09-834-784-305	Sequence 305, App	773	6	0.5	103	4	US-09-615-192A-398	Sequence 398, App
701	6	0.5	35	4	US-09-834-784-306	Sequence 306, App	774	6	0.5	104	3	US-08-535-681-44	Sequence 44, App
702	6	0.5	37	2	US-08-631-328-45	Sequence 45, App	775	6	0.5	104	3	US-08-535-681-48	Sequence 48, App
703	6	0.5	37	2	US-08-631-328-48	Sequence 48, App	776	6	0.5	105	4	US-08-311-731A-227	Sequence 227, App
704	6	0.5	37	2	US-08-631-328-52	Sequence 52, App	777	6	0.5	106	6	5204258-4	Sequence 4307, App
705	6	0.5	37	2	US-08-631-328-53	Sequence 53, App	778	6	0.5	108	4	US-09-107-532A-4307	Sequence 4307, App
706	6	0.5	37	2	US-09-732-210-510	Sequence 510, App	779	6	0.5	110	4	US-09-732-210-299	Sequence 299, App
707	6	0.5	38	1	US-07-977-630-59	Sequence 69, App	780	6	0.5	112	4	US-09-134-001C-3352	Sequence 3352, App
708	6	0.5	45	3	US-08-776-059-11	Sequence 11, App	781	6	0.5	113	4	US-09-378-088A-92	Sequence 92, App
709	6	0.5	45	3	US-09-205-258-554	Sequence 554, App	782	6	0.5	113	4	US-09-378-088A-94	Sequence 94, App
710	6	0.5	46	4	US-09-369-247-127	Sequence 127, App	783	6	0.5	113	4	US-09-107-532A-4336	Sequence 4336, App
711	6	0.5	48	1	US-08-446-692-41	Sequence 41, App	784	6	0.5	114	3	US-08-535-681-42	Sequence 42, App
712	6	0.5	48	2	US-08-488-351A-41	Sequence 41, App	785	6	0.5	115	4	US-08-535-681-50	Sequence 50, App
713	6	0.5	49	2	US-08-248-839C-81	Sequence 81, App	786	6	0.5	115	1	US-08-152-922A-7	Sequence 7, App
714	6	0.5	51	1	US-08-188-228-32	Sequence 32, App	787	6	0.5	115	1	US-09-219-983A-4	Sequence 4, App
715	6	0.5	51	1	US-08-332-643-32	Sequence 32, App	788	6	0.5	117	4	US-09-378-088A-96	Sequence 96, App
716	6	0.5	51	1	US-08-332-638-32	Sequence 32, App	789	6	0.5	117	4	US-09-378-088A-98	Sequence 98, App
717	6	0.5	51	2	US-08-760-075A-27	Sequence 27, App	790	6	0.5	117	4	US-09-378-088A-100	Sequence 100, App
718	6	0.5	51	3	US-09-338-546-37	Sequence 37, App	791	6	0.5	117	4	US-09-378-088A-102	Sequence 102, App
719	6	0.5	51	4	US-09-659-084-27	Sequence 27, App	792	6	0.5	117	4	US-09-378-088A-104	Sequence 104, App
720	6	0.5	52	4	5198345-14	Patent No. 5198345	793	6	0.5	117	4	US-09-378-088A-106	Sequence 106, App
721	6	0.5	64	3	US-09-188-930-122	Sequence 122, App	794	6	0.5	119	3	US-09-023-082A-80	Sequence 80, App
722	6	0.5	64	4	US-09-312-283C-122	Sequence 122, App	795	6	0.5	120	1	US-09-107-532A-5730	Sequence 3700, App
723	6	0.5	64	4	US-09-107-532A-4650	Sequence 4650, App	796	6	0.5	120	1	US-08-249-013-9	Sequence 9, App
724	6	0.5	65	4	US-09-230-041-28	Sequence 28, App	797	6	0.5	120	2	US-08-886-863-9	Sequence 9, App
725	6	0.5	67	4	US-09-230-041-26	Sequence 26, App	798	6	0.5	120	4	US-09-175-229-9	Sequence 9, App
726	6	0.5	69	2	US-08-292-968-21	Sequence 21, App	799	6	0.5	120	5	PCT-US95-06764-9	Sequence 9, App
727	6	0.5	69	2	US-08-467-974-21	Sequence 21, App	800	6	0.5	121	4	US-09-107-532A-7107	Sequence 7107, App
728	6	0.5	69	2	US-08-467-536-21	Sequence 21, App	801	6	0.5	122	4	US-09-107-532A-5730	Sequence 5730, App
729	6	0.5	69	3	US-08-467-976-21	Sequence 21, App	802	6	0.5	122	4	US-09-107-532A-6744	Sequence 6744, App
730	6	0.5	69	3	US-08-995-159-3	Sequence 3, App	803	6	0.5	123	1	US-08-131-625B-17	Sequence 17, App
731	6	0.5	69	3	US-09-082-514-21	Sequence 21, App	804	6	0.5	123	2	US-08-789-464A-13	Sequence 13, App
732	6	0.5	69	4	US-09-134-001C-3787	Sequence 3787, App	805	6	0.5	123	3	US-08-666-568C-11	Sequence 11, App
733	6	0.5	69	4	US-09-545-605-3	Sequence 3, App	806	6	0.5	123	3	US-09-113-750A-42	Sequence 42, App
734	6	0.5	69	4	US-09-732-210-1003	Sequence 1003, App	807	6	0.5	123	4	US-09-331-924-2	Sequence 2, App
735	6	0.5	69	4	US-09-732-210-1004	Sequence 1004, App	808	6	0.5	123	4	US-09-732-210-209	Sequence 209, App
736	6	0.5	70	4	US-09-134-001C-4143	Sequence 4143, App	809	6	0.5	123	5	PCT-US95-09927-13	Sequence 13, App
737	6	0.5	70	4	5204097-1	Patent No. 5204097	810	6	0.5	123	5	PCT-US95-10904-19	Sequence 19, App
738	6	0.5	71	4	US-09-134-001C-4271	Sequence 4271, App	811	6	0.5	123	5	PCT-US95-10904-53	Sequence 53, App
739	6	0.5	73	3	US-09-100-804-28	Sequence 28, App	812	6	0.5	123	5	PCT-US95-10904-55	Sequence 55, App
740	6	0.5	75	3	US-08-480-640A-191	Sequence 191, App	813	6	0.5	123	5	PCT-US95-10904-57	Sequence 57, App
741	6	0.5	75	3	US-08-686-968C-191	Sequence 191, App	814	6	0.5	123	5	PCT-US95-10904-59	Sequence 59, App
742	6	0.5	75	3	US-08-488-237A-191	Sequence 191, App	815	6	0.5	123	5	PCT-US95-10904-61	Sequence 61, App
743	6	0.5	75	3	US-08-375-992A-191	Sequence 191, App	816	6	0.5	126	4	US-08-328-352-7948	Sequence 7949, App
744	6	0.5	75	4	US-08-472-679H-191	Sequence 191, App	817	6	0.5	127	3	US-09-134-001C-5283	Sequence 5283, App
745	6	0.5	76	4	US-09-328-352-4506	Sequence 4506, App	818	6	0.5	129	3	US-08-942-886-3	Sequence 3, App
746	6	0.5	77	4	US-09-328-352-4362	Sequence 4362, App	819	6	0.5	129	3	US-08-942-886-3	Sequence 3, App
747	6	0.5	77	2	US-08-530-280-16	Sequence 16, App	820	6	0.5	130	4	US-09-134-001C-3231	Sequence 3331, App
748	6	0.5	79	4	US-08-858-207A-305	Sequence 305, App	821	6	0.5	132	4	US-09-563-269-4	Sequence 4, App
749	6	0.5	79	4	US-09-328-352-4679	Sequence 4679, App	822	6	0.5	132	4	US-09-732-210-123	Sequence 123, App
750	6	0.5	81	4	US-09-134-001C-5248	Sequence 5248, App	823	6	0.5	134	1	US-08-564-458-1	Sequence 1, App
751	6	0.5	81	4	US-09-198-452A-1186	Sequence 1186, App	824	6	0.5	134	1	US-08-564-833-1	Sequence 1, App
752	6	0.5	86	4	US-09-134-001C-3986	Sequence 3986, App	825	6	0.5	134	1	US-08-535-682-1	Sequence 1, App
753	6	0.5	88	4	US-09-134-001C-4055	Sequence 4055, App	826	6	0.5	134	1	US-08-446-383A-1	Sequence 1, App
754	6	0.5	89	3	US-08-81C-977-4	Sequence 4, App	827	6	0.5	134	1	US-08-446-383A-2	Sequence 2, App
755	6	0.5	89	3	US-09-107-532A-6533	Sequence 6533, App	828	6	0.5	134	1	US-08-795-628-1	Sequence 1, App
756	6	0.5	93	3	US-08-931-858E-51	Sequence 51, App	829	6	0.5	134	1	US-08-519-777-76	Sequence 76, App
757	6	0.5	93	3	US-08-931-858E-134	Sequence 134, App	830	6	0.5	134	1	US-08-519-777-77	Sequence 77, App

831	6	0.5	134	1	US-08-519-777-78	Sequence 78, Appl	904	6	0.5	160	4	US-09-732-210-1604	Sequence 1604, Ap
832	6	0.5	134	1	US-08-618-543-1	Sequence 1, Appl1	905	6	0.5	161	4	US-09-252-991A-30294	Sequence 30294, A
833	6	0.5	134	1	US-08-742-035-76	Sequence 76, Appl	906	6	0.5	162	4	US-09-198-452A-948	Sequence 948, App
834	6	0.5	134	1	US-08-742-035-77	Sequence 77, Appl	907	6	0.5	164	4	US-09-134-001C-5325	Sequence 5325, Ap
835	6	0.5	134	1	US-08-742-035-78	Sequence 78, Appl	908	6	0.5	164	4	US-09-252-991A-30166	Sequence 30166, A
836	6	0.5	134	2	US-08-777-019-76	Sequence 76, Appl	909	6	0.5	164	4	US-09-252-991A-32517	Sequence 32517, A
837	6	0.5	134	2	US-08-777-019-77	Sequence 77, Appl	910	6	0.5	167	2	US-08-747-887-23	Sequence 23, Appl
838	6	0.5	134	2	US-08-777-019-78	Sequence 78, Appl	911	6	0.5	167	2	US-09-328-352-6684	Sequence 6684, Ap
839	6	0.5	134	2	US-08-606-176A-1	Sequence 1, Appl1	912	6	0.5	169	1	US-08-414-926A-12	Sequence 12, Appl
840	6	0.5	134	2	US-08-777-143-76	Sequence 76, Appl	913	6	0.5	169	2	US-08-926-922-12	Sequence 12, Appl
841	6	0.5	134	2	US-08-777-143-77	Sequence 77, Appl	914	6	0.5	169	3	US-09-253-682-12	Sequence 12, Appl
842	6	0.5	134	2	US-08-777-143-78	Sequence 78, Appl	915	6	0.5	169	3	US-09-527-657-12	Sequence 12, Appl
843	6	0.5	134	2	US-08-710-219A-1	Sequence 1, Appl1	916	6	0.5	169	4	US-09-134-001C-4590	Sequence 4590, Ap
844	6	0.5	134	2	US-08-452-242-4	Sequence 4, Appl1	917	6	0.5	169	4	US-09-544-716-17	Sequence 17, Appl
845	6	0.5	134	2	US-08-452-242-6	Sequence 6, Appl1	918	6	0.5	169	4	US-09-557-921-18	Sequence 18, Appl
846	6	0.5	134	3	US-08-453-176A-4	Sequence 4, Appl1	919	6	0.5	171	1	US-08-313-075A-32	Sequence 32, Appl
847	6	0.5	134	3	US-08-453-176A-6	Sequence 6, Appl1	920	6	0.5	171	3	US-09-023-082A-116	Sequence 116, App
848	6	0.5	134	3	US-08-775-414-76	Sequence 76, Appl	921	6	0.5	172	3	US-09-252-991A-28151	Sequence 28151, A
849	6	0.5	134	3	US-08-775-414-77	Sequence 77, Appl	922	6	0.5	174	3	US-09-023-082A-43	Sequence 43, Appl
850	6	0.5	134	3	US-08-775-414-78	Sequence 78, Appl	923	6	0.5	174	3	US-09-023-082A-55	Sequence 55, Appl
851	6	0.5	134	3	US-08-451-374-4	Sequence 4, Appl1	924	6	0.5	174	3	US-09-023-082A-66	Sequence 66, Appl
852	6	0.5	134	3	US-08-451-374-6	Sequence 6, Appl1	925	6	0.5	174	3	US-08-941-155-2	Sequence 2, Appl1
853	6	0.5	134	3	US-08-535-681-2	Sequence 2, Appl1	926	6	0.5	174	4	US-08-936-162A-451	Sequence 451, App
854	6	0.5	134	3	US-08-935-268A-4	Sequence 4, Appl1	927	6	0.5	174	4	US-09-872-047-2	Sequence 2, Appl1
855	6	0.5	134	3	US-08-935-268A-6	Sequence 6, Appl1	928	6	0.5	176	4	US-09-252-991A-21613	Sequence 21613, A
856	6	0.5	134	3	US-08-931-858B-76	Sequence 76, Appl	929	6	0.5	178	2	US-08-680-326-4	Sequence 4, Appl1
857	6	0.5	134	3	US-08-931-858B-77	Sequence 77, Appl	930	6	0.5	178	4	US-09-328-352-6697	Sequence 6697, Ap
858	6	0.5	134	3	US-08-931-858B-78	Sequence 78, Appl	931	6	0.5	179	4	US-09-328-352-7612	Sequence 7612, Ap
859	6	0.5	134	3	US-08-981-739-76	Sequence 76, Appl	932	6	0.5	181	1	US-08-185-432-6	Sequence 6, Appl1
860	6	0.5	134	3	US-08-981-739-77	Sequence 77, Appl	933	6	0.5	182	1	US-09-134-001C-5022	Sequence 5022, Ap
861	6	0.5	134	3	US-08-981-739-78	Sequence 78, Appl	934	6	0.5	182	4	US-09-328-352-6638	Sequence 6638, Ap
862	6	0.5	134	3	US-09-220-528-13	Sequence 13, Appl	935	6	0.5	184	3	US-09-023-082A-30	Sequence 30, Appl
863	6	0.5	134	4	US-08-452-229-4	Sequence 4, Appl1	936	6	0.5	184	3	US-08-741-411-1	Sequence 1, Appl1
864	6	0.5	134	4	US-08-452-229-6	Sequence 6, Appl1	937	6	0.5	185	1	US-07-741-940-6	Sequence 6, Appl1
865	6	0.5	134	4	US-09-128-026-76	Sequence 76, Appl1	938	6	0.5	185	1	US-08-289-546A-6	Sequence 6, Appl1
866	6	0.5	134	4	US-09-128-026-77	Sequence 77, Appl	939	6	0.5	185	1	US-08-452-654-6	Sequence 6, Appl1
867	6	0.5	134	4	US-09-128-026-78	Sequence 78, Appl	940	6	0.5	185	1	US-08-452-655B-6	Sequence 6, Appl1
868	6	0.5	134	4	US-09-732-210-1566	Sequence 1566, Ap	941	6	0.5	185	2	US-08-865-338-4	Sequence 4, Appl1
869	6	0.5	135	3	US-08-535-681-47	Sequence 47, Appl	942	6	0.5	185	3	US-08-450-582-6	Sequence 6, Appl1
870	6	0.5	136	3	US-08-905-223-464	Sequence 464, App	943	6	0.5	185	4	US-08-449-731-6	Sequence 6, Appl1
871	6	0.5	136	4	US-09-107-532A-6435	Sequence 6435, Ap	944	6	0.5	186	4	US-09-996-243-211	Sequence 211, App
872	6	0.5	137	4	US-09-452-991A-27555	Sequence 27555, A	945	6	0.5	186	4	US-09-252-991A-30615	Sequence 30615, A
873	6	0.5	138	4	US-09-180-422B-12	Sequence 12, Appl	946	6	0.5	186	4	US-09-107-532A-1526	Sequence 1526, Ap
874	6	0.5	138	4	US-09-107-532A-4755	Sequence 4755, Ap	947	6	0.5	189	4	US-09-107-532A-4526	Sequence 4526, Ap
875	6	0.5	138	6	5218098-1	Patent No. 5218098	948	6	0.5	191	4	US-09-252-991A-29829	Sequence 29829, A
876	6	0.5	139	3	US-09-367-953B-116	Sequence 116, App	949	6	0.5	193	4	US-08-671-545C-40	Sequence 40, Appl
877	6	0.5	139	4	US-09-352-991A-21299	Sequence 21299, A	950	6	0.5	194	4	US-09-134-001C-3563	Sequence 3563, Ap
878	6	0.5	140	4	US-09-328-352-7601	Sequence 7601, Ap	951	6	0.5	194	4	US-09-149-476-735	Sequence 735, App
879	6	0.5	140	4	US-09-732-210-1255	Sequence 1255, Ap	952	6	0.5	194	4	US-09-252-991A-21846	Sequence 21846, A
880	6	0.5	143	4	US-09-134-001C-4711	Sequence 4711, Ap	953	6	0.5	195	4	US-09-487-792-11	Sequence 11, Appl
881	6	0.5	144	4	US-09-134-001C-3102	Sequence 3102, Ap	954	6	0.5	195	4	US-09-908-594-11	Sequence 11, Appl
882	6	0.5	144	4	US-09-352-991A-26540	Sequence 26540, A	955	6	0.5	196	4	US-09-252-991A-24314	Sequence 24314, A
883	6	0.5	145	4	US-09-352-991A-22630	Sequence 22630, A	956	6	0.5	197	4	US-09-328-352-5000	Sequence 5000, Ap
884	6	0.5	147	4	US-09-134-001C-5392	Sequence 5392, A	957	6	0.5	199	3	US-09-103-359-1	Sequence 1, Appl1
885	6	0.5	149	4	US-09-198-452A-792	Sequence 792, App	958	6	0.5	200	3	US-09-113-750A-40	Sequence 40, Appl
886	6	0.5	151	4	US-09-328-352-6429	Sequence 6429, Ap	959	6	0.5	200	4	US-09-562-737-101	Sequence 101, App
887	6	0.5	152	4	US-09-732-210-53	Sequence 53, Appl	960	6	0.5	200	4	US-09-562-737-102	Sequence 102, App
888	6	0.5	154	4	US-09-252-991A-17864	Sequence 17864, A	961	6	0.5	200	4	US-09-562-737-109	Sequence 109, App
889	6	0.5	154	4	US-09-198-452A-239	Sequence 239, App	962	6	0.5	200	4	US-09-562-737-110	Sequence 110, App
890	6	0.5	155	2	US-08-844-057-4	Sequence 4, Appl1	963	6	0.5	200	4	US-09-252-991A-27855	Sequence 27855, A
891	6	0.5	155	2	US-09-006-730-4	Sequence 4, Appl1	964	6	0.5	204	2	US-08-760-075A-16	Sequence 16, Appl
892	6	0.5	155	4	US-09-107-532A-5729	Sequence 5729, Ap	965	6	0.5	204	3	US-09-338-546-16	Sequence 16, Appl
893	6	0.5	157	2	US-08-630-822A-60	Sequence 60, Appl	966	6	0.5	204	4	US-09-659-084-16	Sequence 16, Appl
894	6	0.5	157	2	US-09-005-069-60	Sequence 60, Appl	967	6	0.5	205	4	US-09-252-991A-18239	Sequence 18239, A
895	6	0.5	157	4	US-09-171-156A-18	Sequence 18, Appl	968	6	0.5	205	4	US-09-198-452A-829	Sequence 829, App
896	6	0.5	157	4	US-09-004-730A-18	Sequence 18, Appl	969	6	0.5	206	4	US-09-134-001C-5047	Sequence 5047, Ap
897	6	0.5	157	4	US-09-252-991A-22957	Sequence 22957, A	970	6	0.5	207	4	US-09-252-991A-19505	Sequence 29505, A
898	6	0.5	157	4	US-08-981-799A-18	Sequence 18, Appl	971	6	0.5	207	4	US-09-434-354-40	Sequence 40, Appl
899	6	0.5	158	4	US-09-433-248A-4	Sequence 4, Appl1	972	6	0.5	208	1	US-08-462-169B-20	Sequence 20, Appl
900	6	0.5	159	4	US-09-252-991A-16672	Sequence 16672, A	973	6	0.5	208	2	US-08-951-822-30	Sequence 30, Appl
901	6	0.5	159	4	US-09-352-991A-20158	Sequence 20158, A	974	6	0.5	208	3	US-09-103-079-20	Sequence 20, Appl
902	6	0.5	159	4	US-09-452-991A-21183	Sequence 21183, Ap	975	6	0.5	208	3	US-09-023-082A-2	Sequence 2, Appl1
903	6	0.5	160	4	US-09-328-352-6929	Sequence 6929, Ap	976	6	0.5	208	3	US-09-023-082A-30	Sequence 20, Appl


```
977 6 0.5 208 3 US-09-023-082A-24 Sequence 24, Appl
978 6 0.5 208 3 US-09-023-082A-39 Sequence 39, Appl
979 6 0.5 208 3 US-09-218-44A-2 Sequence 2, Appl
980 6 0.5 208 4 US-09-358-951-30 Sequence 30, Appl
981 6 0.5 208 4 US-09-425-021-20 Sequence 20, Appl
982 6 0.5 208 4 US-09-229-947-30 Sequence 30, Appl
983 6 0.5 208 4 US-09-252-991A-21959 Sequence 21959, A
984 6 0.5 209 4 US-09-914-999-14 Sequence 14, Appl
985 6 0.5 209 4 US-09-198-452A-286 Sequence 286, Appl
986 6 0.5 210 1 US-07-741-940-4 Sequence 4, Appl
987 6 0.5 210 1 US-08-289-548A-4 Sequence 4, Appl
988 6 0.5 210 1 US-08-452-654-4 Sequence 4, Appl
989 6 0.5 210 1 US-08-452-655B-4 Sequence 4, Appl
990 6 0.5 210 3 US-08-450-582-4 Sequence 4, Appl
991 6 0.5 210 4 US-09-134-001C-3671 Sequence 3671, Ap
992 6 0.5 210 4 US-08-449-731-4 Sequence 4, Appl
993 6 0.5 217 4 US-08-679-493A-196 Sequence 196, App
994 6 0.5 218 4 US-09-252-991A-22136 Sequence 22136, A
995 6 0.5 220 1 US-07-820-154A-4 Sequence 4, Appl
996 6 0.5 220 2 US-08-097-554A-4 Sequence 4, Appl
997 6 0.5 220 3 US-08-453-176A-26 Sequence 26, Appl
998 6 0.5 220 3 US-08-480-640A-4 Sequence 4, Appl
999 6 0.5 220 3 US-08-451-374-26 Sequence 26, Appl
1000 6 0.5 220 3 US-08-295-802-4 Sequence 4, Appl
```

ALIGNMENTS

```
RESULT 1
US-09-384-162-6
; Sequence 6, Application US/09384162
; Patent No. 6376747
```

```
; GENERAL INFORMATION:
; APPLICANT: King, Ti
; APPLICANT: Malik, Kamal
; APPLICANT: Martin-Heller, Teresa
; APPLICANT: Miki L., Brian
; TITLE OF INVENTION: No. 6376747e1 Plant-Derived Map Kinase Kinase
; FILE REFERENCE: 08-884280US
; CURRENT APPLICATION NUMBER: US/09/384,162
; CURRENT FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Dictyostelium discoideum
; US-09-384-162-6
```

```
Query Match 1.2%; Score 16; DB 4; Length 286;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 742 NNNNNNNNNNNNNNNNN 757
Db 206 NNNNNNNNNNNNNNNNN 221
```

```
RESULT 2
US-09-718-841-4
; Sequence 4, Application US/09718841
; Patent No. 6333184
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6333184e1 motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1051
; CURRENT APPLICATION NUMBER: US/09/718,841
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 4
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Human
; NAME/KEY: VARIANT
; LOCATION: (16)...(37)
; OTHER INFORMATION: Xaa = any amino acid
US-09-718-841-4
```

```
Query Match 1.2%; Score 16; DB 4; Length 357;
Best Local Similarity 100.0%; Pred. No. 3.2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 92 FAYGQTSGKTYTMLG 107
Db 98 FAYGQTSGKTYTMLG 113
```

```
RESULT 3
US-09-718-810-4
; Sequence 4, Application US/09718810
; Patent No. 6420162
```

```
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6420162e1 motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1051
; CURRENT APPLICATION NUMBER: US/09/718,810
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Human
; NAME/KEY: VARIANT
; LOCATION: (16)...(37)
; OTHER INFORMATION: Xaa = any amino acid
US-09-718-810-4
```

```
Query Match 1.2%; Score 16; DB 4; Length 357;
Best Local Similarity 100.0%; Pred. No. 3.2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 92 FAYGQTSGKTYTMLG 107
Db 98 FAYGQTSGKTYTMLG 113
```

```
RESULT 4
US-09-118-319-5
; Sequence 5, Application US/09118319
; Patent No. 6114158
```

```
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Chen, Huizhong
; APPLICANT: Ljungdahl, Lars G.
; TITLE OF INVENTION: Orythomyces Cellulase Celf Protein and Coding Sequences
; FILE REFERENCE: 33-98sequence listing
; CURRENT APPLICATION NUMBER: US/09/118,319
; CURRENT FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Neocallimastix patriciarum
; US-09-118-319-5
```

```
Query Match 1.2%; Score 16; DB 3; Length 428;
```


Best Local Similarity 100.0%; Pred. No. 3.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNN 757
DB 68 NNNNNNNNNNNNNNNN 83

RESULT 5

US-09-118-319-2
; Sequence 2, Application US/09118319
; Patent No. 6114158
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Chen, Huizhong
; APPLICANT: Ljungdahl, Lars G.
; TITLE OF INVENTION: Oryzomyces Cellulase Celf Protein and Coding Sequences
; FILE REFERENCE: 33-98sequence listing
; CURRENT APPLICATION NUMBER: US/09/118,319
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Oryzomyces sp. PC-2
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
US-09-118-319-2

Query Match 1.2%; Score 16; DB 3; Length 432;
Best Local Similarity 100.0%; Pred. No. 3.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNN 757
DB 78 NNNNNNNNNNNNNNNN 93

RESULT 6

US-09-718-841-2
; Sequence 2, Application US/09718841
; Patent No. 6333184
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6333184el motor proteins and methods for
; FILE REFERENCE: 1051
; CURRENT APPLICATION NUMBER: US/09/718,841
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 563
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (16)...(37)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: VARIANT
; LOCATION: (391)...(412)
; OTHER INFORMATION: Xaa = any amino acid
US-09-718-841-2

Query Match 1.2%; Score 16; DB 4; Length 563;
Best Local Similarity 100.0%; Pred. No. 4.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 FAYGOTSGKTYTMLG 107
DB 98 FAYGOTSGKTYTMLG 113

RESULT 7

US-09-718-810-2
; Sequence 2, Application US/09718810
; Patent No. 6420162
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6420162el motor proteins and methods for
; FILE REFERENCE: 1051
; CURRENT APPLICATION NUMBER: US/09/718,810
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 563
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (16)...(37)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: VARIANT
; LOCATION: (391)...(412)
; OTHER INFORMATION: Xaa = any amino acid
US-09-718-810-2

Query Match 1.2%; Score 16; DB 4; Length 563;
Best Local Similarity 100.0%; Pred. No. 4.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 FAYGOTSGKTYTMLG 107
DB 98 FAYGOTSGKTYTMLG 113

RESULT 8

US-08-914-999-8
; Sequence 8, Application US/08914999
; Patent No. 6346406
; GENERAL INFORMATION:
; APPLICANT: Ryzanov, Alexey G.
; APPLICANT: Hailt, William N.
; APPLICANT: Pavut, Karen S.
; TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
; TITLE OF INVENTION: AND METHODS OF USE THEREFOR
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/914,999
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 601-1-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 732 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Dictyostelium discoideum
US-08-914-999-8

Query Match 1.2%; Score 16; DB 4; Length 732;
Best Local Similarity 100.0%; Pred. No. 6,1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNN 757
Db 355 NNNNNNNNNNNNNNNN 370

RESULT 9
US-08-340-011-5
Sequence 5, Application US/08340011
Patent No. 576755
GENERAL INFORMATION:
APPLICANT: Alltalo, et al.
TITLE OF INVENTION: FLT4, A NOVEL RECEPTOR TYROSIN KINASE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
City: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,011
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/959,951
FILING DATE: 09-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Thomas C.
REGISTRATION NUMBER: 36,989
REFERENCE/DOCKET NUMBER: 32267
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1311 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-340-011-5

Query Match 1.2%; Score 16; DB 1; Length 1311;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNN 757
Db 613 NNNNNNNNNNNNNNNN 628

RESULT 10
US-08-901-710-5
Sequence 5, Application US/089017710
Patent No. 6107046
GENERAL INFORMATION:
APPLICANT: Alltalo, Kari
APPLICANT: Aprelikova, Olga
APPLICANT: Pajusola, Katri
APPLICANT: Armstrong, Eilina
APPLICANT: Koivonen, Jaana
APPLICANT: Kaipainen, Arja
APPLICANT: Matikainen, Marja-Terttu
TITLE OF INVENTION: FLT4, A RECEPTOR TYROSINE KINASE, AND USES
THEREOF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
City: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/901,710
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,011
FILING DATE: 14-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/257,754
FILING DATE: 09-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/959,951
FILING DATE: 09-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Gaas, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28113/33824
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1311 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-901-710-5

Query Match 1.2%; Score 16; DB 3; Length 1311;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 NNNNNNNNNNNNNNNN 757
Db 613 NNNNNNNNNNNNNNNN 628

RESULT 11
US-09-457-040B-27
Sequence 27, Application US/09457040B
Patent No. 6387641
GENERAL INFORMATION:
APPLICANT: Vertex Pharmaceuticals Incorporated


```

? APPLICANT: Bellon, Steve
? TITLE OF INVENTION: Crystallized p38 Complexes
? FILE REFERENCE: VPI/98-14
? CURRENT APPLICATION NUMBER: US/09/457,040B
? CURRENT FILING DATE: 1999-12-08
? NUMBER OF SEQ ID NOS: 41
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 27
? LENGTH: 1584
? TYPE: prt
? ORGANISM: DICDI - Dictyostelium Discoideum
? US-09-457-040B-27

```

Query Match	1.2%	Score 16;	DB 4;	Length 1584;
Best Local Similarity	100.0%	Pred. No. 1.2e-06;		
Matches 16;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	742	NNNNNNNNNNNNNNNN	757
Db	449	NNNNNNNNNNNNNNNN	464

```

RESULT 12
US-09-417-485D-6
: Sequence 6, Application US/09417485D
: Patent No. 6541202
: GENERAL INFORMATION:
: APPLICANT: Long, David M.
: APPLICANT: Metz, Anneke M.
: APPLICANT: Love, Ruschelle A.
: TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes
: FILE REFERENCE: 4/774-5009-US
: CURRENT APPLICATION NUMBER: US/09/417,485D
: CURRENT FILING DATE: 2002-06-14
: NUMBER OF SEQ ID NOS: 49
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 6
: LENGTH: 2184
: TYPE: PRT
: ORGANISM: Plasmodium falciparum
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (330)..(335)
: OTHER INFORMATION: xaa at position 330 = leu or ile;
: OTHER INFORMATION: xaa at position 335 = asp or gly.
: US-09-417-485D-6

```

Query Match	1.2%	Score 16	DB 4	Length 2184
Best Local Similarity	100.0%	Pred. No. 1.6e-06		
Matches 16	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Qy	742	NNNNNNNNNNNNNNNNNNNN	757
Db	239	NNNNNNNNNNNNNNNNNNNN	254

```

RESULT 13
US-09-722-129-4
: Sequence 4, Application US/09722129
: Patent No. 6582958
: GENERAL INFORMATION:
: APPLICANT: Beraud, Christophe
: APPLICANT: Freedman, Richard
: TITLE OF INVENTION: No. 6582958e1 motor proteins and methods for
: TITLE OF INVENTION: their use
: FILE REFERENCE: 1054
: CURRENT APPLICATION NUMBER: US/09/722,129
: CURRENT FILING DATE: 2000-11-24
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 303
: TYPE: PRT
:

```

US-09-722-129-4

Query Match	1.2%;	Score 15;	DB 4;	Length 303;
Best Local Similarity	100.0%;	Pred. No. 2.6e-06;		
Matches 15;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      91 CFAYGQTSGSKTYTM 105
          |||||
Db      90 CFAYGQTSGSKTYTM 104
```

```

RESULT 14
US-09-722-129-2
; Sequence 2, Application US/09722129
; Patent No. 6582958
; GENERAL INFORMATION:
; APPLICANT: Berard, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6582958e1 motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1054
; CURRENT APPLICATION NUMBER: US/09/7722.129
; CURRENT FILING DATE: 2000-11-24
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Human
; US-09-722-129-2

```

Query Match	1.2%;	Score 15;	DB 4;	Length 492;
Best Local Similarity	100.0%;	Pred. No. 4.1e-06;		
Matches 15;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	91	CFAYGQTSGSKTYTM	105
Db	279	CFAYGQTSGSKTYTM	293

```

RESULT 15
US-09-417-485D-8
: Sequence 8, Application US/09417485D
: Patent No. 6541202
: GENERAL INFORMATION:
: APPLICANT: Long, David M.
: APPLICANT: Metz, Anneke M.
: APPLICANT: Love, Ruschelle A.
: TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes
: FILE REFERENCE: 47714-5009-US
: CURRENT APPLICATION NUMBER: US/09/417,485D
: CURRENT FILING DATE: 2002-06-14
: NUMBER OF SEQ ID NOS: 49
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 8
: LENGTH: 794
: TYPE: PRT
: ORGANISM: Plasmodium falciparum
US-09-417-485D-8

```

Query Match 1.2%; Score 15; DB 4; Length 794;
Best Local Similarity 100.0%; Pred. No. 6.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	744	NNNNNNNNNNNNNNNI	758
D6	130	NNNNNNNNNNNNNNNI	144

RESULT 16
US-08-861-464-6
; Sequence 6, Application US/08861464

Patent No. 5874210
GENERAL INFORMATION:
APPLICANT: Guarente, Leonard P.
APPLICANT: Austriaco Jr., Nicanor
TITLE OF INVENTION: Genes Determining Cellular Senescence
TITLE OF INVENTION: In Yeast
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
ZIP: 02173
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/861,464
FILING DATE: 22-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/396,001
FILING DATE: 28-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09351
FILING DATE: 15-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,408
FILING DATE: 16-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-6408A2Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: 781-861-6240
TELEFAX: 781-861-9540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 888 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-861-464-6
Query Match 1.2%; Score 15; DB 2; Length 888;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 NNNNNNNNNNNNNN 756
Db 521 NNNNNNNNNNNNNN 535
RESULT 17
US-08-396-001-6
Sequence 6, Application US/08396001
Patent No. 5919618
GENERAL INFORMATION:
APPLICANT: Guarente, Leonard P.
APPLICANT: Austriaco Jr., Nicanor
APPLICANT: Cole, James
APPLICANT: Cole, Francesca
APPLICANT: Kennedy, Brian
TITLE OF INVENTION: Genes Determining Cellular Senescence in
TITLE OF INVENTION: Yeast
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive

CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/396,001
FILING DATE: 28-FEB-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-6408A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-9540
TELEFAX: 617-861-6240
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 888 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-396-001-6
Query Match 1.2%; Score 15; DB 2; Length 888;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 NNNNNNNNNNNNNN 756
Db 521 NNNNNNNNNNNNNN 535
RESULT 18
US-09-323-433A-6
Sequence 6, Application US/09323433A
Patent No. 6218512
GENERAL INFORMATION:
APPLICANT: Guarente, Leonard P.
APPLICANT: Austriaco Jr., Nicanor
APPLICANT: Cole, James
APPLICANT: Cole, Francesca
APPLICANT: Kennedy, Brian
TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESCENCE IN
TITLE OF INVENTION: YEAST
FILE REFERENCE: 0050,1491-003
CURRENT APPLICATION NUMBER: US/09/323,433A
CURRENT FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: US 08/396,001
PRIOR FILING DATE: 1995-02-28
PRIOR APPLICATION NUMBER: PCT/US94/09351
PRIOR FILING DATE: 1994-08-15
PRIOR APPLICATION NUMBER: US 08/107,408
PRIOR FILING DATE: 1993-08-16
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 888
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-323-433A-6
Query Match 1.2%; Score 15; DB 3; Length 888;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 NNNNNNNNNNNNNN 756
Db 521 NNNNNNNNNNNNNN 535


```

RESULT 19
US-09-098-901-11
; Sequence 11, Application US/09098901B
; Patent No. 6218144
; GENERAL INFORMATION:
; APPLICANT: Scott, Matthew
; APPLICANT: Sisson, John C.
; TITLE OF INVENTION: Costal2 Genes and their Uses
; FILE REFERENCE: SUN-65P
; CURRENT APPLICATION NUMBER: US/09/098,901B
; CURRENT FILING DATE: 1998-06-17
; EARLIER APPLICATION NUMBER: 60/051,347
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 23
; TYPE: PRT
; ORGANISM: CONSENSUS
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(23)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-098-901-11

```

```

Query Match
Best Local Similarity 1.1%; Score 14; DB 3; Length 23;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 92 FAYGOTSGKTYTM 105
Db 8 FAYGOTSGKTYTM 21

```

```

RESULT 20
US-09-641-806-4
; Sequence 4, Application US/09641806
; Patent No. 6395527
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6395527e1 motor proteins and methods for
; FILE REFERENCE: 1034
; CURRENT APPLICATION NUMBER: US/09/641,806
; CURRENT FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Human
US-09-641-806-4

```

```

Query Match
Best Local Similarity 1.1%; Score 14; DB 4; Length 337;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 92 FAYGOTSGKTYTM 105
Db 80 FAYGOTSGKTYTM 93

```

```

RESULT 21
US-09-723-129-4
; Sequence 4, Application US/09723129
; Patent No. 6551787
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6551787e1 motor proteins and methods for

```

```

; TITLE OF INVENTION: their use
; FILE REFERENCE: 1034
; CURRENT APPLICATION NUMBER: US/09/723,129
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/641,806
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Human
US-09-723-129-4

```

```

Query Match
Best Local Similarity 1.1%; Score 14; DB 4; Length 337;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 92 FAYGOTSGKTYTM 105
Db 80 FAYGOTSGKTYTM 93

```

```

RESULT 22
US-09-722-862-4
; Sequence 4, Application US/09722862
; Patent No. 6562610
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6562610e1 motor proteins and methods for
; FILE REFERENCE: 1034
; CURRENT APPLICATION NUMBER: US/09/722,862
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/641,806
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Human
US-09-722-862-4

```

```

Query Match
Best Local Similarity 1.1%; Score 14; DB 4; Length 337;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 92 FAYGOTSGKTYTM 105
Db 80 FAYGOTSGKTYTM 93

```

```

RESULT 23
US-09-641-806-2
; Sequence 2, Application US/09641806
; Patent No. 6395527
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6395527e1 motor proteins and methods for
; FILE REFERENCE: 1034
; CURRENT APPLICATION NUMBER: US/09/641,806
; CURRENT FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Human
US-09-641-806-2

```


Query Match 1.1%; Score 14; DB 4; Length 342;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 92 FAYGOTSGKTYTM 105
Db 81 FAYGOTSGKTYTM 94

RESULT 24
US-09-723-129-2
; Sequence 2, Application US/09723129
; Patent No. 6551787
; GENERAL INFORMATION:
; APPLICANT: Berard, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6551787e1 motor proteins and methods for
; FILE REFERENCE: 1034
; CURRENT APPLICATION NUMBER: US/09/723,129
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/641,806
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Human
US-09-723-129-2

Query Match 1.1%; Score 14; DB 4; Length 342;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 92 FAYGOTSGKTYTM 105
Db 81 FAYGOTSGKTYTM 94

RESULT 25
US-09-722-862-2
; Sequence 2, Application US/09722862
; Patent No. 6562610
; GENERAL INFORMATION:
; APPLICANT: Berard, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6562610e1 motor proteins and methods for
; FILE REFERENCE: 1034
; CURRENT APPLICATION NUMBER: US/09/722,862
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/641,806
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Human
US-09-722-862-2

Query Match 1.1%; Score 14; DB 4; Length 342;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 92 FAYGOTSGKTYTM 105
Db 81 FAYGOTSGKTYTM 94

RESULT 26

US-09-291-170A-3
; Sequence 3, Application US/09291170A
; Patent No. 6410687
; GENERAL INFORMATION:
; APPLICANT: Vale, Ronald D.
; APPLICANT: Hartman, James J.
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Assays for the Detection of Microtubule
; FILE REFERENCE: 18557B-000510US
; CURRENT APPLICATION NUMBER: US/09/291,170A
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 60/081,734
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Xenopus laevis
; OTHER INFORMATION: Xenopus kinesin central motor 1 (XKCM1)
US-09-291-170A-3

Query Match 1.1%; Score 14; DB 4; Length 730;
Best Local Similarity 100.0%; Pred. No. 5.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 263 INRSLALKECIRA 276
Db 520 INRSLALKECIRA 533

RESULT 27
US-09-724-884-3
; Sequence 3, Application US/09724884
; Patent No. 6429304
; GENERAL INFORMATION:
; APPLICANT: Vale, Ronald D.
; APPLICANT: Hartman, James J.
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Assays for the Detection of Microtubule
; FILE REFERENCE: 18557B-000510US
; CURRENT APPLICATION NUMBER: US/09/724,884
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/291,170
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Xenopus laevis
; OTHER INFORMATION: Xenopus kinesin central motor 1 (XKCM1)
US-09-724-884-3

Query Match 1.1%; Score 14; DB 4; Length 730;
Best Local Similarity 100.0%; Pred. No. 5.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 263 INRSLALKECIRA 276
Db 520 INRSLALKECIRA 533

RESULT 28
US-08-006-676B-1
; Sequence 1, Application US/08006676B
; Patent No. 5411865
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven

TITLE OF INVENTION: Diagnosis of leishmaniasis
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jeffrey B. Oster
STREET: 8339 SE 57th Street
CITY: Mercer Island
STATE: Washington
COUNTRY: USA
ZIP: 98040-4906
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/006,676B
FILING DATE: 15-JAN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oster, Jeffrey B.
REGISTRATION NUMBER: 32,585
REFERENCE/DOCKET NUMBER: REED-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 232 7845
TELEFAX: (206) 236 0205
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 955 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-006-676B-1

Query Match 1.1%; Score 14; DB 1; Length 955;
Best Local Similarity 100.0%; Pred. No. 7.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 FAYGOTSGKTYTM 105
|||||
DB 119 FAYGOTSGKTYTM 132

RESULT 29
US-08-282-845-2
Sequence 2, Application US/08282845
Patent No. 5719263
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: A 230kd Antigen Present in Leishmania
TITLE OF INVENTION: Species
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple Macintosh Operating System 7.1
SOFTWARE: Microsoft Word for Macintosh 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/282,845
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/006,676
FILING DATE: JANUARY 15, 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 5004-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 955 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-282-845-2

Query Match 1.1%; Score 14; DB 1; Length 955;
Best Local Similarity 100.0%; Pred. No. 7.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 FAYGOTSGKTYTM 105
|||||
DB 119 FAYGOTSGKTYTM 132

RESULT 30
US-08-428-414A-3
Sequence 3, Application US/08428414A
Patent No. 5912166
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
TITLE OF INVENTION: LEISHMANIASIS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,414A
FILING DATE: 21-APR-1995
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Kadlecik, Ann T.
REGISTRATION NUMBER: 39,244
REFERENCE/DOCKET NUMBER: 210121.407
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836 SEEDANDBERRY
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 955 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-428-414A-3

Query Match 1.1%; Score 14; DB 2; Length 955;
Best Local Similarity 100.0%; Pred. No. 7.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 FAYGOTSGKTYTM 105
|||||
DB 119 FAYGOTSGKTYTM 132

RESULT 31

PCT-US94-00324-1
; Sequence 1, Application PC/TUS9400324
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven
; TITLE OF INVENTION: Diagnosis of Leishmaniasis
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple System 7.1
; SOFTWARE: Microsoft Word, version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00324
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/006,676
; FILING DATE: 15-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 5004-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 955 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-00324-1

Query Match 1.1%; Score 14; DB 5; Length 955;
Best Local Similarity 100.0%; Pred. No. 7.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 92 FAYGQTSGSKTYT 105
Db 119 FAYGQTSGSKTYT 132

RESULT 32
US-09-724-517-4
; Sequence 4, Application US/09724517
; Patent No. 6379941
; GENERAL INFORMATION:
; APPLICANT: Berard, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6379941e1 motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1031
; CURRENT APPLICATION NUMBER: US/09/724,517
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US/09/641,807
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Human
US-09-724-517-4

Query Match 1.0%; Score 13; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 92 FAYGQTSGSKTYT 104
Db 80 FAYGQTSGSKTYT 92

RESULT 33
US-09-641-807A-4
; Sequence 4, Application US/09641807A
; Patent No. 6440731
; GENERAL INFORMATION:
; APPLICANT: Berard, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6440731e1 motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1031
; CURRENT APPLICATION NUMBER: US/09/641,807A
; CURRENT FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Human
US-09-641-807A-4

Query Match 1.0%; Score 13; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 92 FAYGQTSGSKTYT 104
Db 80 FAYGQTSGSKTYT 92

RESULT 34
US-09-723-096-4
; Sequence 4, Application US/09723096
; Patent No. 6448026
; GENERAL INFORMATION:
; APPLICANT: Berard, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6448026e1 motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1031
; CURRENT APPLICATION NUMBER: US/09/723,096
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US/09/641,807
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Human
US-09-723-096-4

Query Match 1.0%; Score 13; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 92 FAYGQTSGSKTYT 104
Db 80 FAYGQTSGSKTYT 92

RESULT 35
US-09-177-165A-31
; Sequence 31, Application US/09177165A
; Patent No. 6426205
; GENERAL INFORMATION:
; APPLICANT: Tyers, Mike


```
APPLICANT: Willems, Andrew
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN
FILE OF INVENTION: DEPENDENT PROTEOLYSIS
FILE REFERENCE: 11757.10USU1
CURRENT APPLICATION NUMBER: US/09/117,165A
CURRENT FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: 60/092,443
PRIOR FILING DATE: 1998-07-10
PRIOR APPLICATION NUMBER: 60/063,254
PRIOR FILING DATE: 1997-10-24
NUMBER OF SEQ ID NOS: 50
SOFTWARE: Patencin Ver. 2.1
SEQ ID NO 31
LENGTH: 1151
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-177-165A-31
```

```
Query Match      1.0%; Score 13; DB 4; Length 1151;
Best Local Similarity 100.0%; Pred. No. 0.0008;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      502 NNNNNNNNNNNNS 514
DB      38 NNNNNNNNNNNNS 50
```

```
RESULT 36
US-09-724-517-2
Sequence 2, Application US/09724517
Patent No. 6379941
GENERAL INFORMATION:
APPLICANT: Beraud, Christophe
APPLICANT: Freedman, Richard
TITLE OF INVENTION: No. 6379941e1 motor proteins and methods for
FILE OF INVENTION: their use
FILE REFERENCE: 1031
CURRENT APPLICATION NUMBER: US/09/724,517
CURRENT FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: US/09/641,807
PRIOR FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1279
TYPE: PRT
ORGANISM: Human
FEATURE:
NAME/KEY: VARIANT
LOCATION: (409)...(436)
OTHER INFORMATION: Xaa = any amino acid
US-09-724-517-2
```

```
Query Match      1.0%; Score 13; DB 4; Length 1279;
Best Local Similarity 100.0%; Pred. No. 0.0008;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      92 PAYGOTGSGKTYT 104
DB      82 PAYGOTGSGKTYT 94
```

```
RESULT 37
US-09-641-807A-2
Sequence 2, Application US/09641807A
Patent No. 6440731
GENERAL INFORMATION:
APPLICANT: Beraud, Christophe
APPLICANT: Freedman, Richard
TITLE OF INVENTION: No. 6440731e1 motor proteins and methods for
FILE OF INVENTION: their use
FILE REFERENCE: 1031
CURRENT APPLICATION NUMBER: US/09/641,807A
```

```
CURRENT FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1279
TYPE: PRT
ORGANISM: Human
FEATURE:
NAME/KEY: VARIANT
LOCATION: (409)...(446)
OTHER INFORMATION: Xaa = any amino acid
US-09-641-807A-2
```

```
Query Match      1.0%; Score 13; DB 4; Length 1279;
Best Local Similarity 100.0%; Pred. No. 0.0008;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      92 PAYGOTGSGKTYT 104
DB      82 PAYGOTGSGKTYT 94
```

```
RESULT 38
US-09-723-096-2
Sequence 2, Application US/09723096
Patent No. 6448026
GENERAL INFORMATION:
APPLICANT: Beraud, Christophe
APPLICANT: Freedman, Richard
TITLE OF INVENTION: No. 6448026e1 motor proteins and methods for
FILE OF INVENTION: their use
FILE REFERENCE: 1031
CURRENT APPLICATION NUMBER: US/09/723,096
CURRENT FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: US/09/641,807
PRIOR FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1279
TYPE: PRT
ORGANISM: Human
FEATURE:
NAME/KEY: VARIANT
LOCATION: (409)...(436)
OTHER INFORMATION: Xaa = any amino acid
US-09-723-096-2
```

```
Query Match      1.0%; Score 13; DB 4; Length 1279;
Best Local Similarity 100.0%; Pred. No. 0.0008;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      92 PAYGOTGSGKTYT 104
DB      82 PAYGOTGSGKTYT 94
```

```
RESULT 39
US-09-451-117-2
Sequence 2, Application US/09451117
Patent No. 6277973
GENERAL INFORMATION:
APPLICANT: Jenkins, Mark C.
APPLICANT: Payer, Ronald
TITLE OF INVENTION: Cloning and Expression of a DNA Sequence Encoding A 41
Patent No. 6277973
TITLE OF INVENTION: kDa Cryptosporidium parvum Oocyst Wall Protein
FILE REFERENCE: 0046.99
CURRENT APPLICATION NUMBER: US/09/451,117
CURRENT FILING DATE: 1999-11-30
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patencin Ver. 2.1
```


; SEQ ID NO 2
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Cryptosporidium parvum
US-09-451-117-2

Query Match 0.9%; Score 12; DB 3; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 741 TNNNNNNNNNN 752
|||||
Db 235 TNNNNNNNNNN 246

RESULT 40
US-09-888-655-2
; Sequence 2, Application US/09888655
; Patent No. 6521229
; GENERAL INFORMATION:
; APPLICANT: Jenkins, Mark C.
; APPLICANT: Fayer, Ronald
; APPLICANT: Trout, James
; TITLE OF INVENTION: Cloning and Expression of a DNA Sequence Encoding A 41
; Patent No. 6521229
; TITLE OF INVENTION: kda Cryptosporidium parvum Oocyst Wall Protein
; FILE REFERENCE: 0046.99
; CURRENT APPLICATION NUMBER: US/09/888,655
; CURRENT FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Cryptosporidium parvum
US-09-888-655-2

Query Match 0.9%; Score 12; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 741 TNNNNNNNNNN 752
|||||
Db 235 TNNNNNNNNNN 246

Search completed: October 2, 2003, 16:36:31
Job time : 36 secs